NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

SIMULATION STUDY OF ESTIMATORS FOR THE SURVIVAL PROBABILTY OF A FIRST PASSAGE TIME FOR A SEMI-MARKOV PROCESS USING CENSORED DATA

by

Park, Byung Goo

September 1988

Thesis Advisor: Patricia A. Jacobs

Unclassified

security classification of this page

REPORT DOCUMENTATION PAGE

1a Report Security Classification Unclassified	1b Restrictive Markings
2a Security Classification Authority	3 Distribution Availability of Report Approved for public release; distribution is unlimited.
2b Declassification Downgrading Schedule	
4 Performing Organization Report Number(s)	5 Monitoring Organization Report Number(s)
6a Name of Performing Organization Naval Postgraduate School — 6b Office Symbol *(if applicable)* 55	7a Name of Monitoring Organization Naval Postgraduate School
6c Address *(city, state, and ZIP code)* Monterey, CA 93943-5000	7b Address *(city, state, and ZIP code)* Monterey, CA 93943-5000
8a Name of Funding Sponsoring Organization — 8b Office Symbol *(if applicable)*	9 Procurement Instrument Identification Number
8c Address *(city, state, and ZIP code)*	10 Source of Funding Numbers Program Element No — Project No — Task No — Work Unit Accession No

11 Title *(include security classification)* SIMULATION STUDY OF ESTIMATORS FOR THE SURVIVAL PROBABILITY OF A FIRST PASSAGE TIME FOR A SEMI-MARKOV PROCESS USING CENSORED DATA

12 Personal Author(s) Park, Byung Goo

13a Type of Report	13b Time Covered	14 Date of Report *(year, month, day)*	15 Page Count
Master's Thesis	From To	September 1988	76

16 Supplementary Notation The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

17 Cosati Codes			18 Subject Terms *(continue on reverse if necessary and identify by block number)*
Field	Group	Subgroup	Semi-Markov, Kaplan-Meier Estimator, Confidence Interval, Jackknife.

19 Abstract *(continue on reverse if necessary and identify by block number)*

Finite state space semi-Markov processes find application in many areas. Often interest centers on whether or not the process has hit a particular state before a time t. This thesis reports results of a simulation study of the small sample behavior for three estimators of the survival probability of a first passage time for a semi-Markov process using censored data.

20 Distribution Availability of Abstract ☒ unclassified unlimited ☐ same as report ☐ DTIC users	21 Abstract Security Classification Unclassified	
22a Name of Responsible Individual Patricia A. Jacobs	22b Telephone *(include Area code)* (408) 646-2258	22c Office Symbol 55jc

DD FORM 1473,84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete — security classification of this page

Unclassified

Approved for public release; distribution is unlimited.

Simulation Study of Estimators for the Survival Probability of a First Passage Time for a Semi-Markov Process Using Censored Data

by

Park, Byung Goo
Major, Republic Of Korea Air Force
B.S., Korea Air Force Academy, 1980

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
September 1988

ABSTRACT

Finite state space semi-Markov processes find application in many areas. Often interest centers on whether or not the process has hit a particular state before a time t. This thesis reports results of a simulation study of the small sample behavior for three estimators of the survival probability of a first passage time for a semi-Markov process using censored data.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACKNOWLEDGEMENTS

For their significant contributions toward this thesis, I would like to thank the following people :

- Professor P. A. Jacobs for her tolerance, professional assistance, and thorough review.
- Professor P. Purdue for his encouragement and guidance.
- Most especially, my wife, Eun-Hye, my son, Sung-Jae, and my daughter, Jee-Eun for their prayers, patience, and earnest support.

I. INTRODUCTION

Finite state space semi-Markov models find applications in a variety of fields such as queueing theory, reliability, and clinical trials. Often interest in the application of these models centers on the distribution of the first passage time to a state or a set of states representing, for example, the lifetime of a system or the end of a busy period of a server. Suppose that the observations of the path of the semi-Markov process are all that is known about the process. The problem is to estimate the probability that the first passage time has not occurred before time t.

Censored data problems arise frequently in medical, and also in engineering system reliability applications. For example, in medical survivorship studies some subjects may be lost to follow-up, or available data may be analyzed before all subjects have expired. In the equipment reliability context, observed units may still be in operation, perhaps after several previous failures, at the time of the analysis.

Three possible estimators will be considered. The three estimators use different amounts of information concerning the process. One estimator uses only the observed first passage times. Another estimator makes parametric assumptions concerning the sojourn time distribution and uses maximum likelihood. A third approach uses an exponential approximation to the probability and empirical distributions to estimate the sojourn time distributions.

The three estimators were investigated for a specific semi-Markov process with uncensored data in Kim[Ref. 1] and Jacobs[Ref. 2]. Results of a simulation study of the three estimators using censored data are reported in Gallagher [Ref. 3]. The emphasis in this latter study is on the behavior of the point estimates with mean biases and standard errors being given.

In this thesis the investigation of the behavior of the three estimators with censored data is continued. The emphasis here is on confidence intervals for the three estimators. In Chapter 2, the three estimators are described and the respective confidence interval procedures considered introduced. Chapter 3 contains the details of the simulation experiment and its results. Finally, conclusions from the study are given in Chapter 4.

II. NATURE OF PROBLEM

A. PROBLEM

The semi-Markov process model considered is as follows. Suppose we observe N individuals. Let $X_t(i)$ be the state of the i^{th} individual at time t. We will assume $\{X_t(i); t \geq 0\}$ $i = 1,2,3,\ldots,N$. are independent identically distributed semi-Markov processes having the same probability law as $\{X_t; t \geq 0\}$. The process $\{X_t; t \geq 0\}$ is a semi-Markov process with three states $\{0, 1, 2\}$. The individuals start at t = 0 in state 1. Upon leaving state 1, the process transitions to state 0 with probability θ and to state 2 with probability $1 - \theta$. From state 2 the process transitions to state 1 with probability 1. State 0 is an absorbing state. The first passage time to state 0 will be referred to as the time of death. The N individuals are censored independently. The censoring times are exponentially distributed with a mean of $\frac{1}{c}$. The entire path of transitions and sojourn times are observed until the time of censoring or death.

Let

$$D = inf\{t \geq 0; X_t = 0\}$$

and

$$S(t) = P\{D > t\}$$

where D is the time of death (or entrance to state 0). The problem is to estimate the survival probability $P\{D > t\}$ with the censored data of N individuals.

B. ESTIMATORS

1. Kaplan-Meier Estimator

A non-parametric estimator of the distribution function for censored data is the Kaplan-Meier estimator(K.M.E.) which is often called the product limit estimator [Ref. 4]. Let $U_1, U_2, \ldots, U_n$ be independent identically distributed random variables with distribution G having a density function. Let $V_1, V_2, \ldots, V_n$ be independent identically distributed times to censoring with a continuous distribution function. Let

$$Z_i = \min(U_i, V_i),$$

and

$$\delta_i = \begin{cases} 0 & \text{if } U_i \le V_i \\ 1 & \text{Otherwise.} \end{cases}$$

The Z_i are the observed times and δ_i is an indicator of whether or not the i^{th} observation is censored. Let $Z_{(1)} \le Z_{(2)} \le \dots \le Z_{(n)}$ be the order statistics of $\{Z_i\}$ and $\delta_{(i)}$ be the corresponding values of $\{\delta_i\}$. It is assumed that there will be no ties since the underlying distribution functions are continuous. The Kaplan-Meier estimate of the survival function $S(t) = \{1 - G(t)\}$ is

$$\hat{S}(t) = \begin{cases} \prod_{[i:Z_{(i)} \le t]} \left[\frac{(n-i)}{(n-i+1)} \right]^{1-\delta_{(i)}} & \text{if } t < Z_{(n)} \\ 0 & \text{if } t > Z_{(n)} \,\&\, \delta_{(n)} = 0 \\ \text{Undefined} & \text{if } t > Z_{(n)} \,\&\, \delta_{(n)} = 1. \end{cases} \tag{2.1}$$

The variance of $\hat{S}(t)$ is given approximately by

$$Var[\hat{S}(t)] \simeq [\hat{S}(t)]^2 \sum_{[i:Z_{(i)} \le t]} \frac{\delta_i}{(n-i)(n-i+1)} \tag{2.2}$$

[Ref. 4: p. 464]. The Kaplan-Meier estimator using the death times of the N individuals will be denoted by $P_K\{D > t\}$.

If the Kaplan-Meier estimate is undefined, we investigate the effect of two methods, defined as MOD 1 and MOD 2, to make the Kaplan-Meier estimate honest. MOD 1 and MOD 2 are defined as follow. In MOD 1, the remaining mass of the estimated survival function is assigned to the last datum $Z_{(n)}$(which is censored). In MOD 2, if the last k data points are censored the remaining mass of the survival function is distributed equally among the k data points. For example, if the estimated survival probability at the last uncensored point is 0.2 and there are additional two data points which are censored, then each of these additional points is assigned a mass of $\frac{0.2}{2}$.

2. Maximum Likelihood Estimator

In this subsection, the maximum likelihood estimator(M.L.E.) will be given for the special case in which the sojourn time in state i is exponentially distributed with mean $\frac{1}{\rho_i}$ $(i = 1, 2)$.

Let R_{ij} be the number of transitions from state i to state j for one individual. The log likelihood function for an individual is

$$l = R_{12} \ln(1 - \theta) + R_{10} \ln \theta + R_{21} \ln \rho_2 + (R_{10} + R_{12}) \ln \rho_1 - \rho_1 T_1 - \rho_2 T_2 \qquad (2.3)$$

where T_i $(i = 1,2)$ is the total time spent in state i before death or censoring.

The maximum likelihood estimators using the data from all N individuals are

$$\hat{\theta} = \frac{\tilde{R}_{10}}{\tilde{R}_{10} + \tilde{R}_{12}}; \qquad (2.4)$$

$$\hat{\rho}_1 = \frac{\tilde{R}_{10} + \tilde{R}_{12}}{\tilde{T}_1}; \qquad (2.5)$$

$$\hat{\rho}_2 = \frac{\tilde{R}_{21}}{\tilde{T}_2}; \qquad (2.6)$$

$$\hat{c} = \frac{\#\text{ death times that are censored}}{\tilde{T}_1 + \tilde{T}_2}, \qquad (2.7)$$

where

$$\tilde{R}_{ij} = \sum_{n=1}^{N} R_{ij}(n) \qquad (2.8)$$

and

$$\tilde{T}_i = \sum_{n=1}^{N} T_i(n) \qquad (2.9)$$

with $R_{ij}(n)$ being the number of transitions from i to j for the n^{th} individual and $T_i(n)$ being the total time spent in state i before death or censoring for the i^{th} individual.

To obtain asymptotic variances for these estimators, note that

$$\frac{-\partial^2 l}{\partial\theta^2} = \frac{R_{10}}{\theta^2} + \frac{R_{12}}{(1-\theta)^2}\,; \tag{2.10}$$

$$\frac{-\partial^2 l}{\partial\rho_1^2} = \frac{R_{10}+R_{12}}{\rho_1^2}\,; \tag{2.11}$$

$$\frac{-\partial^2 l}{\partial\rho_2^2} = \frac{R_{21}}{\rho_2^2}\,. \tag{2.12}$$

Further

$$E[R_{12}] = \sum_{n=1}^{\infty} P\{R_{12} \geq n\} \tag{2.13}$$

$$= \sum_{n=1}^{\infty} (1-\theta)\frac{\rho_1}{\rho_1+c}\,[(\frac{\rho_1}{\rho_1+c})(\frac{\rho_2}{\rho_2+c})(1-\theta)]^n$$

$$= \frac{(1-\theta)\rho_1(\rho_2+c)}{\theta\rho_1\rho_2 + c[\rho_1+\rho_2] + c^2}\,;$$

$$E[R_{21}] = \sum_{n=1}^{\infty} P\{R_{21} \geq n\} \tag{2.14}$$

$$= \sum_{n=1}^{\infty} [(1-\theta)\frac{\rho_1}{\rho_1+c}\,\frac{\rho_2}{\rho_2+c}]^n$$

$$= \frac{(1-\theta)\rho_1\rho_2}{\theta\rho_1\rho_2 + c[\rho_1+\rho_2] + c^2}\,;$$

$$E[R_{10}] = \sum_{n=0}^{\infty} \frac{\rho_1}{\rho_1+c}\,\theta[(1-\theta)\{\frac{\rho_1\rho_2}{(\rho_1+c)(\rho_2+c)}\}]^n \tag{2.15}$$

$$= \frac{\rho_1\theta(\rho_2+c)}{\theta\rho_1\rho_2 + c[\rho_1+\rho_2] + c^2}$$

where c^{-1} is the mean of the exponential censoring time. Thus for N individuals the asymptotic variances of the estimators are

$$Var[\hat{\theta}] = I(\theta)^{-1};$$

$$Var[\hat{\rho}_1] = I(\rho_1)^{-1}; \qquad (2.16)$$

$$Var[\hat{\rho}_2] = I(\rho_2)^{-1}$$

where

$$I(\theta) = N\ E[\frac{R_{10}}{\theta^2} + \frac{R_{12}}{(1-\theta)^2}]; \qquad (2.17)$$

$$I(\rho_1) = N\ E[\frac{R_{10} + R_{12}}{\rho_2^2}]; \qquad (2.18)$$

$$I(\rho_2) = N\ E[\frac{R_{21}}{\rho_2^2}]. \qquad (2.19)$$

The expression for the survival function $S(t) = P\{D > t\}$ for this continuous time Markov chain is

$$S(t) = \{\frac{\lambda_2 + \rho_2}{\lambda_2} e^{\lambda_2 t} - \frac{\lambda_1 + \rho_2}{\lambda_1} e^{\lambda_1 t}\}\theta \frac{\rho_1}{\lambda_1 - \lambda_2} \qquad (2.20)$$

where λ_1, λ_2 are the roots of the equation

$$\theta\rho_1\rho_2 + y(\rho_1 + \rho_2) + y^2 = 0. \qquad (2.21)$$

The maximum likelihood estimator, denoted as $\hat{P}_M\{D > t\}$, for the survival probability is [Ref. 5: p. 5 eqn 1.17]

$$\hat{P}_M\{D > t\} = \frac{\hat{\theta}\hat{\rho}_1}{\hat{\lambda}_1 - \hat{\lambda}_2}\left\{\frac{\hat{\lambda}_2 + \hat{\rho}_2}{\hat{\lambda}_2} e^{\hat{\lambda}_2 t} - \frac{\hat{\lambda}_1 + \hat{\rho}_2}{\hat{\lambda}_1} e^{\hat{\lambda}_1 t}\right\} \qquad (2.22)$$

where $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are roots of the equation

$$\hat{\theta}\hat{\rho}_1\hat{\rho}_2 + y(\hat{\rho}_1 + \hat{\rho}_2) + y^2 = 0. \qquad (2.23)$$

Since the maximum likelihood estimators are orthogonal, the asymptotic variance of $\hat{P}_M\{D > t\}$ [Ref. 5: p. 5 eqn 1.19] is approximately

$$Var[\hat{P}_M\{D > t\} | \theta, \rho_1, \rho_2]$$

$$= Var(\hat{\theta})(\frac{\partial S}{\partial \theta})^2 + Var(\hat{\rho}_1)(\frac{\partial S}{\partial \hat{\rho}_1})^2 + Var(\hat{\rho}_2)(\frac{\partial S}{\partial \hat{\rho}_2})^2 \tag{2.24}$$

where

$$\frac{\partial S}{\partial \theta} = \frac{S}{\theta} + \frac{\partial S}{\partial \lambda_1}\frac{\partial \lambda_1}{\partial \theta} + \frac{\partial S}{\partial \lambda_2}\frac{\partial \lambda_2}{\partial \theta}; \tag{2.25}$$

$$\frac{\partial S}{\partial \rho_1} = \frac{S}{\rho_1} + \frac{\partial S}{\partial \lambda_1}\frac{\partial \lambda_1}{\partial \rho_1} + \frac{\partial S}{\partial \lambda_2}\frac{\partial \lambda_2}{\partial \rho_1}; \tag{2.26}$$

$$\frac{\partial S}{\partial \rho_2} = \frac{\theta \rho_1}{(\lambda_1 - \lambda_2)}[\frac{1}{\lambda_2}e^{\lambda_2 t} - \frac{1}{\lambda_1}e^{\lambda_1 t}] + \frac{\partial S}{\partial \lambda_1}\frac{\partial \lambda_1}{\partial_{\rho 2}} + \frac{\partial S}{\partial \lambda_2}\frac{\partial \lambda_2}{\partial \rho_2}; \tag{2.27}$$

$$\frac{\partial S}{\partial \lambda_1} = \frac{-S}{(\lambda_1 - \lambda_2)} + \frac{\theta \rho_1}{(\lambda_1 - \lambda_2)}[\frac{\rho_2}{\lambda_1^2}e^{\lambda_1 t} - (\frac{\rho_2}{\lambda_1} + 1)te^{\lambda_1 t}]; \tag{2.28}$$

$$\frac{\partial S}{\partial \lambda_2} = \frac{S}{(\lambda_1 - \lambda_2)} + \frac{\theta \rho_1}{(\lambda_1 - \lambda_2)}[-\frac{\rho_2}{\lambda_2^2}e^{\lambda_2 t} + (\frac{\rho_2}{\lambda_2} + 1)te^{\lambda_2 t}]; \tag{2.29}$$

$$\frac{\partial \lambda_i}{\partial \theta} = \frac{-\rho_1 \rho_2}{\rho_1 + \rho_2 + 2\lambda_i}; \tag{2.30}$$

$$\frac{\partial \lambda_i}{\partial \rho_1} = \frac{-\lambda_i - \theta \rho_2}{\rho_1 + \rho_2 + 2\lambda_i}; \tag{2.31}$$

$$\frac{\partial \lambda_i}{\partial \rho_2} = \frac{-\lambda_i - \theta \rho_1}{\rho_1 + \rho_2 + 2\lambda_i}. \tag{2.32}$$

3. Asymptotic Renewal Estimator

In this subsection, we describe an asymptotic renewal estimator(A.R.E.) for $P\{D > t\}$. A conditioning argument yields the following equation for $P\{D > t\}$;

$$P\{D > t\} = g(t) + (1-\theta)\int_0^t (F_1 * F_2)(ds)\, P\{D > t - s\} \tag{2.33}$$

with

$$g(t) = \overline{F}_1(t) + (1-\theta)\int_0^t F_1(ds)\, \overline{F}_2(t-s). \tag{2.34}$$

Thus, $P\{D > t\}$ satisfies a renewal-type equation with defective inter-renewal distribution

$$L(t) = (1-\theta)\,(F_1 * F_2)(t) \tag{2.35}$$

where F_i is the sojourn time distribution in state i, $\overline{F}_i(t) = 1 - F_i(t)$, and $(F_1 * F_2)(t)$ denotes the convolution of F_1 and F_2. Following Feller [Ref. 6], let k be such that

$$\int_0^\infty e^{kt}\, L(dt) = (1-\theta)\,\phi_1(k)\,\phi_2(k) = 1 \tag{2.36}$$

where

$$\phi_i(\xi) = \int_0^\infty e^{\xi t}\, F_i(dt). \tag{2.37}$$

Then, under certain integrability conditions, if $(F_1 * F_2)$ is not arithmetic

$$\lim_{t\to\infty} e^{kt}\, P\{D > t\} = \frac{b}{\mu} \tag{2.38}$$

where

$$\mu = \int_0^\infty s\, e^{ks}\, L(ds) \tag{2.39}$$

and

$$b = \int_0^{\infty} e^{ks}\, g(s) ds \tag{2.40}$$

$$= \int_0^{\infty} e^{ks}\, [\bar{F}_1(s) + (1-\theta) \int_0^{\infty} F_1\,(du)\, \bar{F}_2(s-u)] ds$$

$$= \frac{1}{k}\, [\{\phi_1(k) - 1\} + (1-\theta)\, \phi_1(k)\{\phi_2(k) - 1\}].$$

Since

$$(1-\theta)\phi_1(k)\, \phi_2(k) = 1$$

it follows that

$$b = \frac{1}{k}\, [\phi_1(k) - 1] + \frac{1}{\phi_2(k)\, k}\, [\phi_2(k) - 1]$$

$$= \frac{1}{k}\, [\phi_1(k) - 1] + \frac{1}{k}\, [1 - \frac{1}{\phi_2(k)}\,]$$

$$= \frac{1}{k}\, [\phi_1(k) - 1] + \frac{1}{k}\, [1 - (1-\theta)\, \phi_1(k)]$$

$$= \frac{1}{k}\, [\phi_1(k) - 1] + \frac{1}{k}\, [1 - \phi_1(k) + \theta\, \phi_1(k)]$$

$$= \frac{\theta}{k}\, \phi_1(k). \tag{2.41}$$

Let $\hat{F}_i$ be the Kaplan-Meier estimate of F_i, and $\hat{\theta}$ be the maximum likelihood estimate of θ [Ref. 5: p. 10]; put

$$\hat{\phi}_i(\xi) \;=\; \int_0^{\infty} e^{\xi s} \hat{F}_i(ds). \tag{2.42}$$

The asymptotic renewal estimator ($\hat{P}_A\{D > t\}$) [Ref. 5: p. 11 eqn 3.11] of the survival probability $P\{D > t\}$ is

$$\hat{P}_A\{D > t\} = e^{\hat{k}t}\frac{\hat{b}}{\hat{\mu}} \tag{2.43}$$

where $\hat{k}$ is the solution to the equation

$$(1 - \hat{\theta})\hat{\phi}_1(k)\hat{\phi}_2(k) = 1; \tag{2.44}$$

$$\hat{\mu} = (1 - \hat{\theta})\int_0^{\infty} e^{\hat{k}s}\, s\,(\hat{F}_1 * \hat{F}_2)\,(ds); \tag{2.45}$$

$$\hat{b} = \frac{\hat{\theta}}{\hat{k}}\hat{\phi}_1(\hat{k}). \tag{2.46}$$

In the simulation $\hat{k}$ is obtained by using the golden section search method.

The asymptotic renewal estimator is undefined if all the sojourn times for a particular state are censored since the Kaplan-Meier estimator is not defined in this case.

C. CONFIDENCE INTERVAL PROCEDURES

A confidence interval for an unknown parameter gives both an indication of the numerical value of the unknown parameter and a measure of how confident we are of that numerical value. Two statistics L and U form a $(1 - \alpha)100\%$ confidence interval for θ, if under repeated random sampling $L \leq \theta \leq U$ $(1 - \alpha)100\%$ of the time. A confidence interval procedure for an unknown parameter θ is also used to make a decision concerning θ, as in classical hypothesis testing or decision making, or to indicate the accuracy and variability of a point estimator $\hat{\theta}$.

Confidence interval procedures for the three estimators for $P\{D > t\}$ will be described in this section, starting with the confidence interval procedure for the Kaplan-Meier estimator(K.M.E.). Procedures for the maximum likelihood estimator(M.L.E.) and the asymptotic renewal estimator will then be discussed.

A preliminary transformation to approximately symmetrize the sampling distribution of the estimator is often beneficial [Ref. 7]. For this study we consider two transformations, the arc-sine and log transformations, which tend to stabilize and also

approximately symmetrize the data. These transformations were suggested by the work of Gaver and Miller [Ref. 8]. Since the individuals are independent, the number of individuals surviving a fixed time t would have a binomial distribution if there were no censoring. The logarithmic and arc-sine transformations have been beneficial in this case.

The confidence intervals for the Kaplan-Meier estimator and the maximum likelihood estimator are the asymptotic normal confidence intervals using the transformed estimator.

The confidence interval for the arc-sine transformed estimator is computed as follows. Let $\hat{S}(t)$ be the estimator of $S(t) = P\{D > t\}$. Since

$$\frac{d}{dx}\sin^{-1}\sqrt{x} = \frac{1}{2}\frac{1}{\sqrt{1-x}}\frac{1}{\sqrt{x}}, \tag{2.47}$$

a Taylor expansion yields

$$\sin^{-1}(\sqrt{\hat{S}(t)}\,) \simeq \sin^{-1}(\sqrt{S(t)}\,) + \frac{1}{2}\frac{1}{\sqrt{(1-S(t))S(t)}}(\hat{S}(t)-S(t)). \tag{2.48}$$

Hence, the approximate variance of $\sin^{-1}(\sqrt{\hat{S}(t)}\,)$ is

$$Var[\sin^{-1}(\sqrt{\hat{S}(t)}\,)] \simeq \frac{1}{4}\frac{1}{(1-\hat{S}(t))\hat{S}(t)} Var[\hat{S}(t)]. \tag{2.49}$$

If $\hat{S}(t) = 1$, we set $Var[\sin^{-1}(\sqrt{\hat{S}(t)}\,)]$ equal to 0. Therefore, a $(1-\alpha)100\%$ confidence interval for $\sin^{-1}\sqrt{S(t)}$ is

$$\sin^{-1}(\sqrt{\hat{S}(t)}\,) \pm [z_{1-\alpha/2}]\sqrt{Var[\sin^{-1}(\sqrt{\hat{S}(t)}\,)]} \tag{2.50}$$

where $z_{1-\alpha/2}$ is the $(1-\alpha/2)100\%$ point of a standard normal.

A confidence interval for the log transformed estimator of S(t) is computed as follows. Since

$$\frac{d}{dx}\ln x = \frac{1}{x}, \tag{2.51}$$

a Taylor expansion yields

$$ln\,\hat{S}(t) \simeq ln\,S(t) + \frac{1}{S(t)}(\hat{S}(t)-S(t)). \tag{2.52}$$

Thus the approximate variance of ln $S(t)$ is

$$Var[ln\,\hat{S}(t)] \simeq \frac{1}{\hat{S}(t)^2} Var[\hat{S}(t)]. \tag{2.53}$$

If $\hat{S}(t) = 1$, we set $Var[\hat{S}(t)] = 0$. A $(1 - \alpha)100\%$ confidence interval for $ln\,\hat{S}(t)$ is

$$ln\,\hat{S}(t) \;\pm\; (z_{1-\alpha/2})\sqrt{Var[ln\,\hat{S}(t)]} \tag{2.54}$$

where $z_{1-\alpha/2}$ is the $(1 - \alpha/2)100\%$ point of a standard normal.

1. Asymptotic Normal Confidence Intervals for K.M.E.

The asymptotic variance of the Kaplan-Meier estimator $(\hat{P}_K\{D > t\})$ is given by equation (2.2). The asymptotic normal confidence intervals for $\sin^{-1}\sqrt{P_K\{D > t\}}$ and $\ln P_K\{D > t\}$ are evaluated using equations (2.50) and (2.54) respectively. The corresponding confidence intervals for $P_K\{D > t\}$ are formed by inverse transformation. If the lower limit is less than 0 it is set equal to 0 and if the upper limit is greater than 1 it is set equal to 1.

2. Asymptotic Normal Confidence Intervals for M.L.E.

The asymptotic variance for the maximum likelihood estimator $(\hat{P}_M\{D > t\})$ is given by equation (2.22). The asymptotic normal confidence intervals for $\sin^{-1}(\sqrt{\hat{P}_M\{D > t\}}\,)$ and $\ln P_M\{D > t\}$ can be constructed using equations (2.50) and (2.54) respectively. Confidence intervals for $P_M\{D > t\}$ can be obtained by inverse transformation. If the resulting confidence interval has a lower limit less than 0 it is set equal to 0 and if it has an upper limit greater than 1 it is set equal to 1.

3. Jackknife Procedure for A.R.E.

The jackknife technique was first introduced by Quenouille [Ref. 9] and later utilized by Tukey [Ref. 10] for bias reduction and robust interval estimation. A review can be found in Miller [Ref. 11]. The jackknife is designed to do various jobs fairly well, however it is desirable to avoid (in jackknifing) sampling distributions with (i) abrupt ends and (ii) one or more straggling tails, and it is probably desirable to avoid those that are strongly unsymmetrical [Ref. 12]. Confidence intervals for the asymptotic renewal estimator will be obtained using the jackknife procedure on the arc-sine and log transformed estimates.

The jackknife procedure for the confidence interval of the asymptotic renewal estimator $(\hat{P}_A\{D > t\})$ is implemented as follows.

1. Generate data for N individuals.

2. Compute $\hat{P}_A\{D > t\}$ using all data.

3. Transform $\hat{P}_A\{D > t\}$ into $ln\ \hat{P}_A\{D > t\}$ and $\sin^{-1}\sqrt{\hat{P}_A\{D > t\}}$ which will be denoted as YL_{all}, and YS_{all} respectively. Divide the N individuals into n subgroups such that each subgroup contains $\frac{N}{n}$ individuals.

4. Compute $\hat{P}_A\{D > t\}$ leaving out all data of the i^{th} subgroup and transform it into $ln\ \hat{P}_A\{D > t\}$ and $\sin^{-1}\sqrt{\hat{P}_A\{D > t\}}$ which will be denoted as yl_i, and ys_i respectively.

5. Compute the pseudo-values(denoted as YL_{*i}, and YS_{*i});

$$YL_{*i} = n\,YL_{\text{all}} - (n-1)\,yl_i, \tag{2.55}$$

$$YS_{*i} = n\,YS_{\text{all}} - (n-1)\,ys_i. \tag{2.56}$$

6. Compute the average of pseudo-values which is the jackknifed estimate for the transformed asymptotic renewal estimator(denoted as YL_*, and YS_*);

$$YL_* = \frac{1}{n}[YL_{*1} + \ldots + YL_{*n}], \tag{2.57}$$

$$YS_* = \frac{1}{n}[YS_{*1} + \ldots + YS_{*n}], \tag{2.58}$$

and compute the variance of the average of the pseudo-values(denoted as VL_*^2, and VS_*^2);

$$VL_*^2 = \frac{1}{n(n-1)}\sum_{i=1}^{n}(YL_{*i} - YL_*)^2, \tag{2.59}$$

$$VS_*^2 = \frac{1}{n(n-1)} \sum_{i=1}^{n} (YS_{*i} - YS_*)^2. \tag{2.60}$$

7. Compute the (approximate) two-sided $(1 - \alpha)100\%$ confidence intervals for each of the transformed estimators as follows;

$$YL_* \pm t_{(1-\alpha/2)}(n-1)\sqrt{VL_*^2} \tag{2.61}$$

and

$$YS_* \pm t_{(1-\alpha/2)}(n-1)\sqrt{VS_*^2} \tag{2.62}$$

where $t_{1-\alpha/2}(n-1)$ is the $(1 - \alpha/2)100\%$ point of student's t with n-1 degree of freedom.

8. Confidence intervals for $P\{D > t\}$ are obtained by inverse transformation. If the resulting interval has a lower limit less than 0 it is set equal to 0 and if it has an upper limit greater than 1 it is set equal to 1.

III. ANALYSIS OF SIMULATION RESULTS FOR CONFIDENCE INTERVAL PROCEDURES

A. SIMULATION

All simulations were carried out on an IBM 3033AP computer at the Naval Postgraduate School using the LLRANDOM II random number generator package[Ref. 13]. Plots of simulated estimates and confidence intervals were produced by an experimental APL package GRAFSTAT which the Naval Postgraduate School is using under a test agreement with IBM Watson Research Center, Yorktown, Heights, NY.

The data for the simulation experiments are generated as follows;

1. An individual starts in state 1 at time 0.

2. An exponential censoring time with mean 1/c is generated.

3. An exponential sojourn time in state 1 with mean $1/\rho_1$ is generated.

4. The sojourn time and censoring time are compared; if the sojourn time is smaller, then the sojourn time is recorded and given an uncensored index '0'; if the censoring time is smaller, the sojourn time truncated at the censoring time is recorded and given a censored index '1'. In the latter case, the death time is recorded as the truncated sojourn time and associated with a censored index of '1' and the simulation for the first individual is completed.

5. If the process is not censored in state 1, a uniform random number is generated and compared with θ; if less than θ the process jumps to state 0, and the uncensored death time with index '0' is recorded and the simulation for the first individual is completed; if greater than θ, the process jumps to state 2.

6. An exponential sojourn time for state 2 with mean $1/\rho_2$ is generated and the total time(sojourn time in state 1 plus sojourn time in state 2) is compared to the censoring time with the same actions as listed above.

7. If the process is not censored by the end of the sojourn time in state 2, the process jumps to state 1 and continues until an uncensored or censored death occurs. The time is recorded and the next individual is started.

This procedure continues until N individuals' data have been generated. Using this data, the Kaplan-Meier estimate ($\hat{P}_K(t)$), maximum likelihood estimate ($\hat{P}_M(t)$), asymptotic renewal estimate ($\hat{P}_A(t)$), and their respective 90% and 80% two-sided confidence intervals are computed. This completes the one super-replication.The simulation is replicated for SR = 500 super-replications utilizing different seeds to generate the data. For the simulated model described above parameter values of $\rho_1 = 1$, $\rho_2 = 1$, $\theta = 0.5$, and $c = \frac{1}{2}$, $\frac{1}{5}$, $\frac{1}{10}$, and $\frac{1}{100}$ are used, and for each super-replication N = 50 individuals' data are generated.

B. ANALYSIS

In this section results from the simulation experiments will be reported. In Appendix A, the true survival probability, which is obtained using equation (2.20), is given at the various values of t. The times considered are t = 1.0, 3.0, 5.0, 7.0, 9.0, 11.0, 13.0, and 15.0.

Tables 2 through 5 of Appendix B report the number of super-replications that still have defined Kaplan-Meier estimates at various times t. The column headed by KM0 reports the number of super-replications that have defined Kaplan-Meier estimates at time t for the death time survival function. The columns headed KM1, KM2, repectively show the number of super-replications still defined for the Kaplan-Meier estimate of the survival function at time t for the sojourn times in state 1, respectively state 2. These numbers indicate the effect the censoring has on the Kaplan-Meier estimate. As expected, increasing the mean time to censoring increases the number of super-replications censored. In all cases the survival functions for the sojourn times in state 1 and 2 are less heavily censored than that of the death times.

Plots for comparing the methods investigated for the Kaplan-Meier estimator are given in Figures 1 through 3. The difference between the mean of the estimated survival probability and the true survival probability ($\hat{S}(t) - P\{D > t\}$), the relative differences

$((\hat{S}(t) - P\{D > t\})/P\{D > t\})$, and the relative root mean square errors ($RMSE/P\{D > t\}$) are respectively plotted in Figures 1 through 3. The numerical results are recorded in Tables 6 through 17. The means, and root mean square errors(RMSE) from SR = 500 super-replications are computed as follows;

$$\bar{S}(t) = \frac{1}{SR}\sum_{i=1}^{SR}\hat{S}_i(t) \tag{3.1}$$

and

$$RMSE = [\frac{1}{SR}\sum_{i=1}^{SR}(\hat{S}_i(t) - S(t))^2]^{\frac{1}{2}} \tag{3.2}$$

where $\hat{S}_i(t)$ is the point estimate of the true value $S(t)$ at time t in the i^{th} super-replication and SR is the number of super-replication.

If a Kaplan-Meier estimate is undefined, its value is taken to be its last defined value. The modifications to the Kaplan-Meier estimate, MOD 1 and MOD 2, are described in Chapter 2. Since the unmodified Kaplan-Meier survival function may not tend to 0 as $t \to \infty$, it tends to overestimate the true survival probability. On the other hand, since the modifications MOD 1 and MOD 2 make the undefined Kaplan-Meier survival function honest, they do well for large t and have less variability than the unmodified K.M.E.. However the modifications appear to bias the estimates for moderate times t. All these methods improve as the mean censoring time decreases. MOD 1 does slightly better than MOD 2.

In Figures 4 through 9, plots are presented for the comparison between the three estimators. The differences, relative differences, and relative RMSE's are computed and plotted in the same manner as before. Figures 4 through 6 show results for the Kaplan-Meier and asymptotic renewal estimators using the unmodified Kaplan-Meier estimator. Figures 7 through 9 show the results obtained when the Kaplan-Meier estimates are modified to be MOD 1 in both the K.M.E. and A.R.E.. The numerical results are recorded in Appendix C. As expected, the maximum likelihood estimator(M.L.E.), which uses the most correct information about the model, tends to have the smallest

relative RMSE's and differences between estimated and true means. In the case of greatest censoring, the M.L.E. shows a slight bias.

The Kaplan-Meier estimator(K.M.E.), which uses the least information about the model, does well for small times even in the greatest censoring case, but it is worse than the others for large times due to the undefined estimates. It has the largest values of mean square error among the three estimators. Therefore the K.M.E. tends to be more variable than the other estimators.

Not surprisingly, the asymptotic renewal estimator(A.R.E.), which uses an asymptotic exponential approximation to the probability, has high relative RMSE's and differences between estimated and true means for small times t. As t increases the relative RMSE's and differences between the means decrease. Tables 6 through 9 indicate that for moderate to large times the A.R.E. has RMSE's that are comparable or less than those of the M.L.E.. The time at which they become comparable is a function of the amount of censoring. If the mean censoring time is larger, the RMSE's become comparable sooner. Tables 6 through 12 indicate that while modifying the Kaplan-Meier estimate of the survival function of the death times improves it for large times t, it creates a bias for moderate times t. Using modified Kaplan-Meier estimators of the survival functions for the sojourn times in states 1 and 2 does not improve the asymptotic renewal estimator.

In Figures 10 through 17, plots are presented for comparing the confidence interval procedures of the three estimators. In order to compare the performance of the confidence interval procedures, we use the following measures: the coverage fraction(CVR), the average half length(AHL) of the confidence interval, and the standard deviation(SHL) of the half lengths. The AHL is determined by summing the half lengths of the confidence intervals of the replications and dividing that sum by the number of replications(SR). The SHL is computed by summing the square of the differences between the individual half lengths and the average half length, dividing that value by (SR-1) and taking the square root of the resulting value. Among confidence interval procedures which achieve the desired coverage rate(0.90, and 0.80), the confidence interval procedure which yields the smallest AHL is preferred. Also preferred is a small SHL representing a stable confidence interval procedure.

For each procedure, the number of intervals covering the true value $P\{D > t\}$ is recorded as well as the number of intervals that are too high or too low. These results are reported in Tables 18, 20, 22, 24, 26, 28, 30, and 32 (Appendix D). Next to each coverage count is given the corresponding coverage fraction in parentheses. If a

$(1-\alpha)100\%$ confidence interval procedure is performing well, then this interval should cover about $(1-\alpha)100\%$ of time. A 95% confidence interval for the coverage fraction is computed using $P \pm 1.96[P(1-P)/SR]^{.5}$,where P is the proportion of $(1-\alpha)\%$ confidence intervals that cover the true value of $P\{D > t\}$ and SR is the number of super-replications. Therefore, if a 80%(90%) confidence interval procedure is working well, then out of 500 super-replications between 382(436) and 418(463) confidence intervals should cover the true value. These intervals correspond to 0.7649(0.8351) and 0.8737(0.9263) respectively. The average half lengths of the confidence intervals for $\hat{S}(t)$ are reported in tables 19, 21, 23, 25, 27, 29, 31, and 33 (Appendix D). The standard deviation of the half length is given in parentheses below the average half length. If an estimator is performing well, its confidence interval should not only have the correct coverage fraction but also a small average half length.

In Figures 10 through 13, the coverage fraction and the average half length of confidence intervals using the log transformation for each estimator are presented. In Figures 14 through 17, the coverage fraction and the average half length of confidence intervals using the arc-sine transformation are presented. The two horizontal lines in the coverage plot show the 95% confidence interval for the coverage fraction.

The following remarks concern the confidence intervals obtained using the arc-sine transformation. The asymptotic normal confidence intervals for the M.L.E. have the correct coverage for small to moderate times but tend to slightly undercover for large times. The asymptotic nomal confidence intervals for the K.M.E. have the correct coverage for small times but undercover for moderate to large times due to the undefined estimates. As the mean censoring time increases, the coverage of the Kaplan-Meier confidence intervals improves. The confidence intervals for the asymptotic renewal estimate tend to undercover. Using a modified Kaplan-Meier estimate makes very little difference in the results.

The following remarks concern the confidence intervals obtained by using the log transformation. Once again the confidence intervals for the K.M.E. undercover for moderate and large times t. The coverage is slightly worse than that obtained using the arc-sine transformation. The confidence intervals for the M.L.E. have the correct coverage and the smallest AHL. The confidence intervals for the Jackknifed A.R.E. have the correct coverage for all but the 80% interval at $t=1$ for $c=1/2$. The average half lengths of the A.R.E. confidence intervals are larger than those of the M.L.E. confidence intervals. Using a modified Kaplan-Meier estimate makes very little difference in the results.

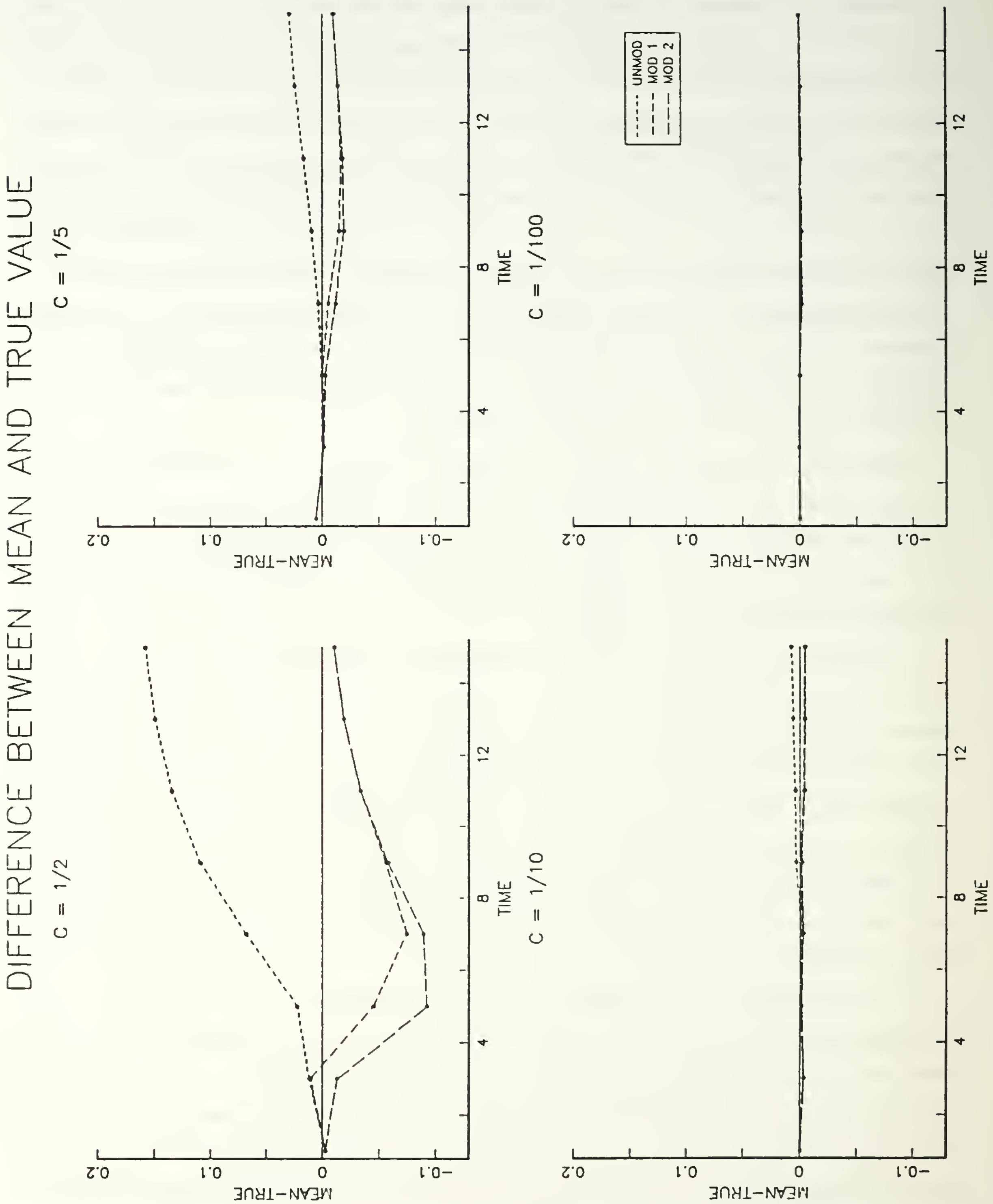

Figure 1. **COMPARISON BETWEEN METHODS FOR K.M.E.:** Difference between Mean and True value.

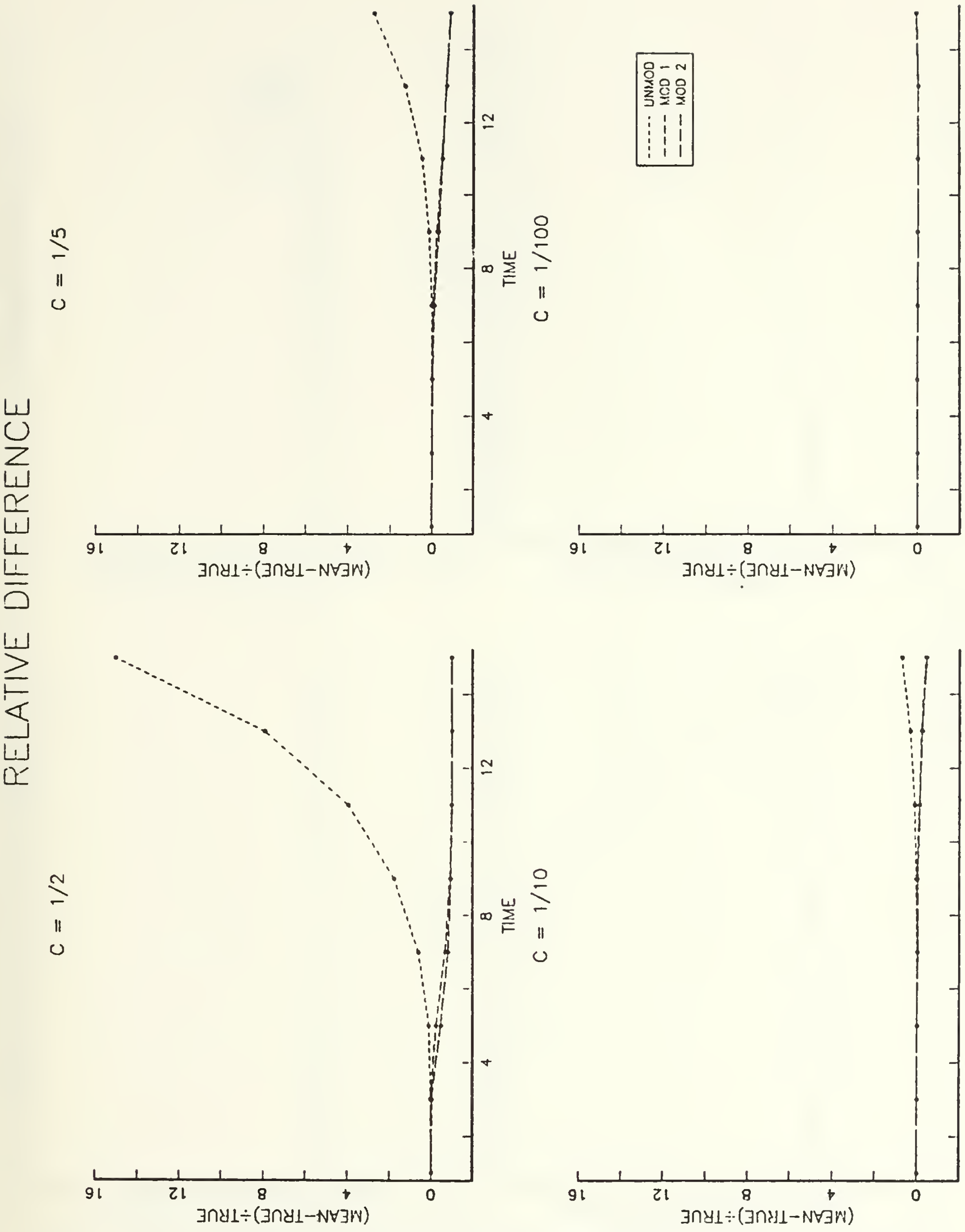

Figure 2. COMPARISON BETWEEN METHODS FOR K.M.E.: Relative Difference between Mean and True value.

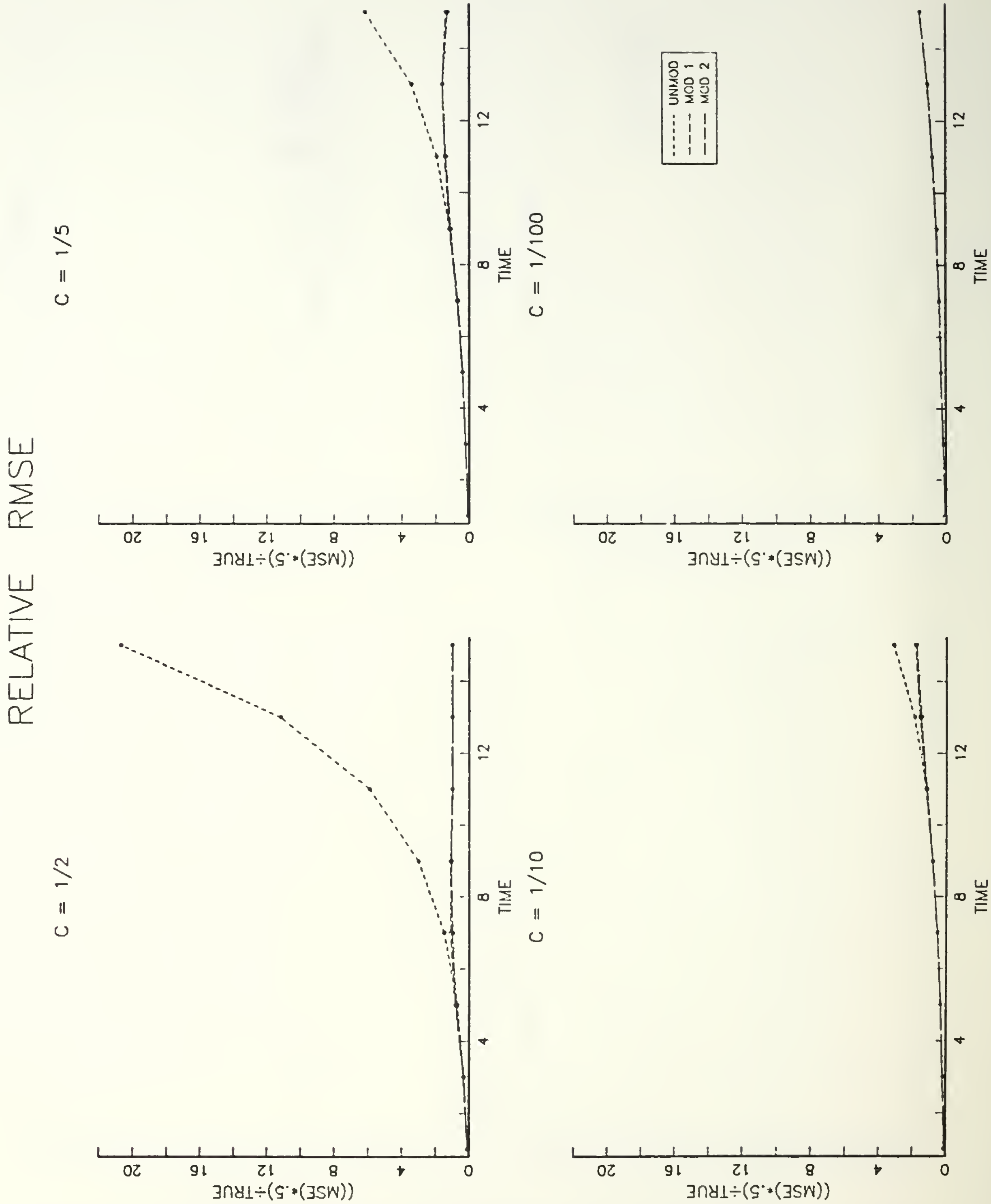

Figure 3. COMPARISON BETWEEN METHODS FOR K.M.E.: Relative Root Mean Square Error.

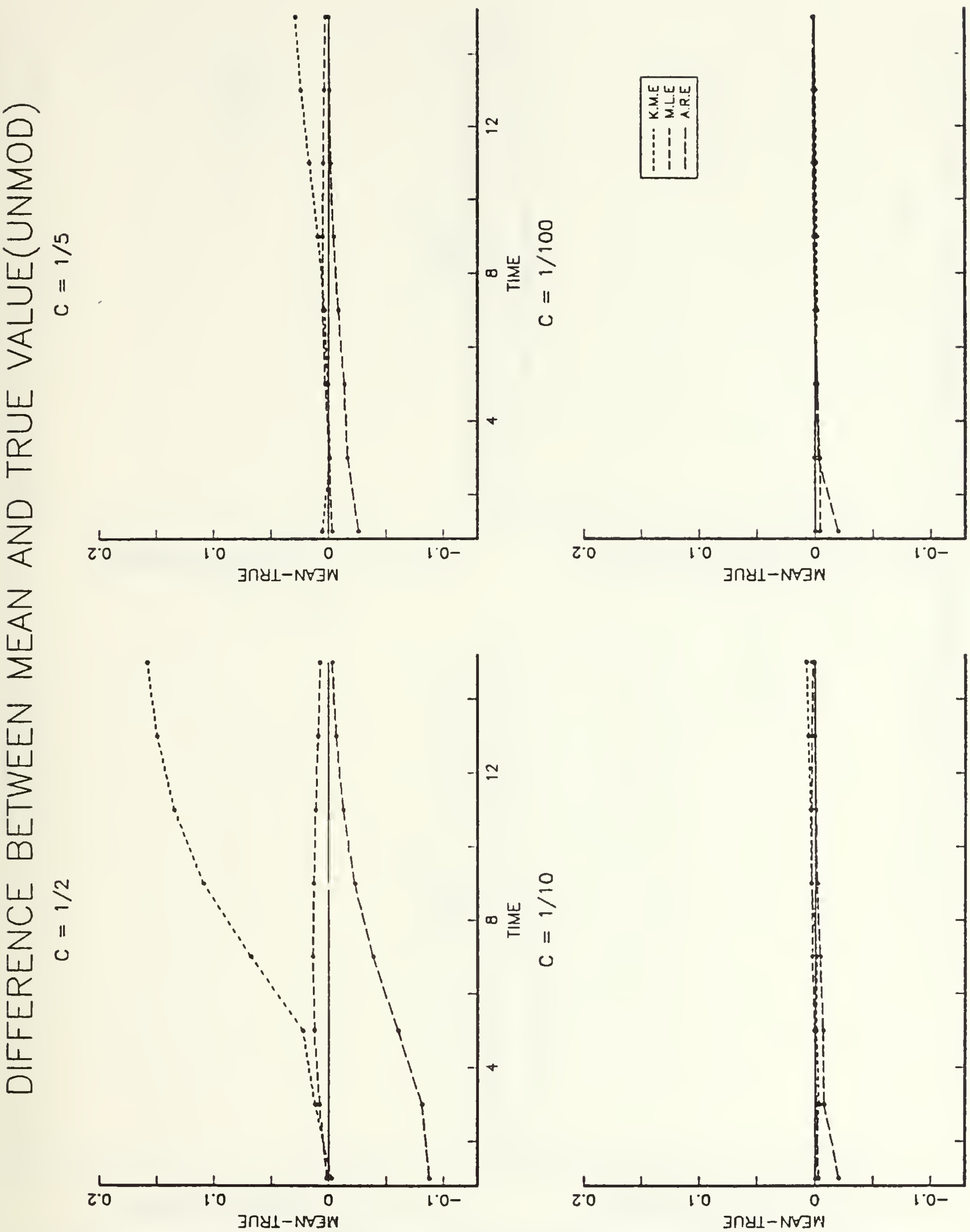

Figure 4. COMPARISON BETWEEN ESTIMATORS(UNMOD): Difference between Mean and True value.

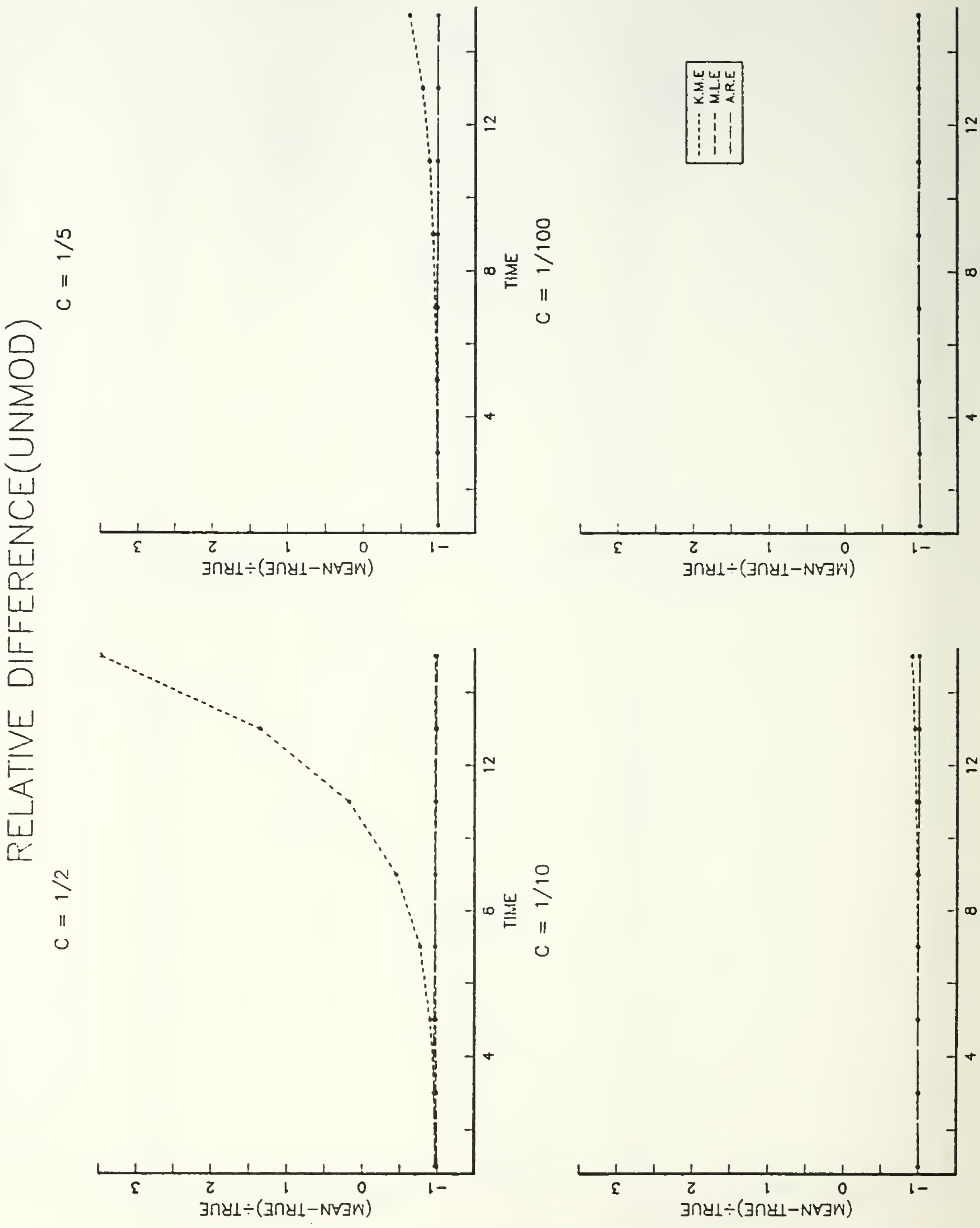

Figure 5. COMPARISON BETWEEN ESTIMATORS(UNMOD): Relative Difference between Mean and True value.

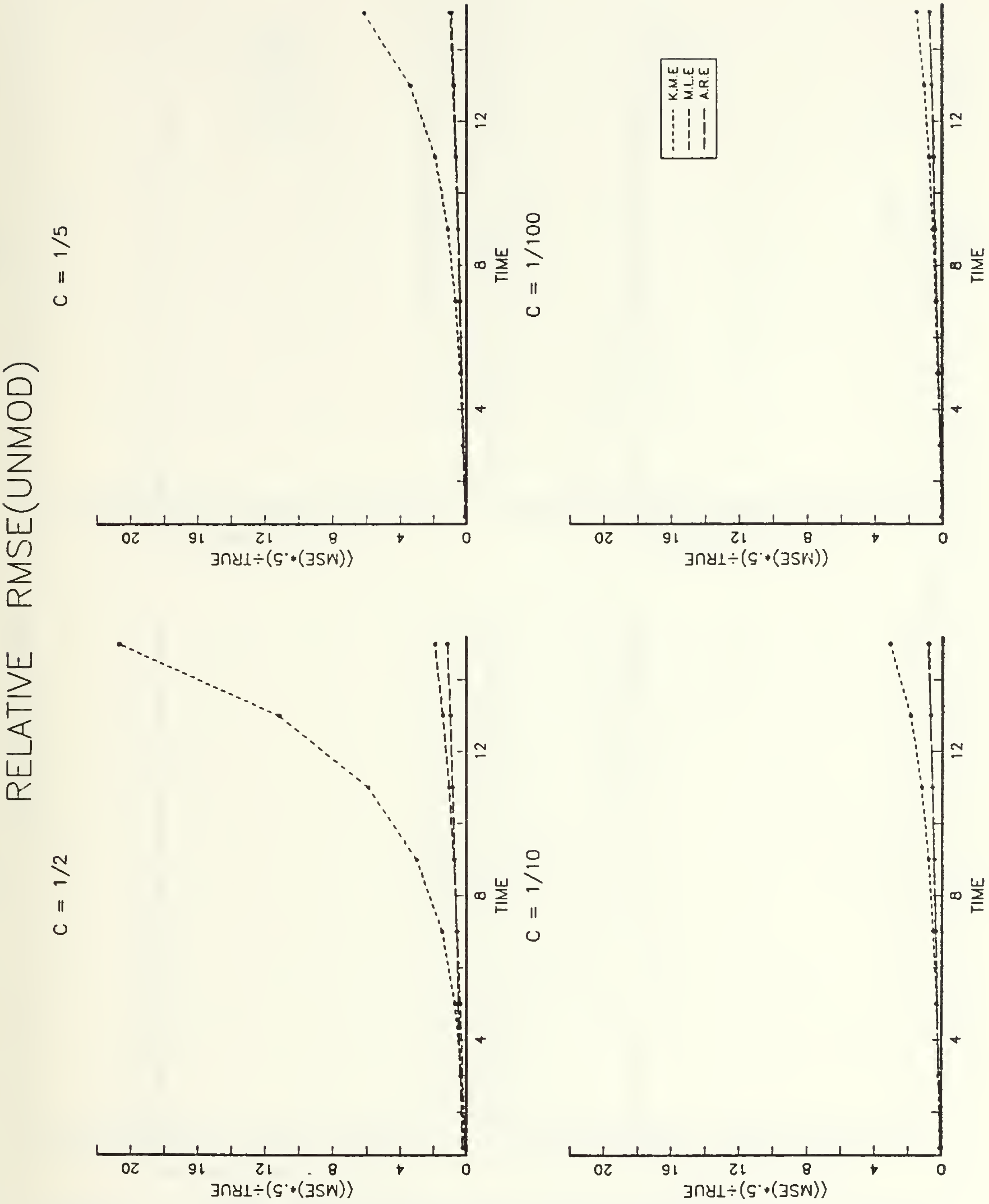

Figure 6. COMPARISON BETWEEN ESTIMATORS(UNMOD): Relative Root Mean Square Error.

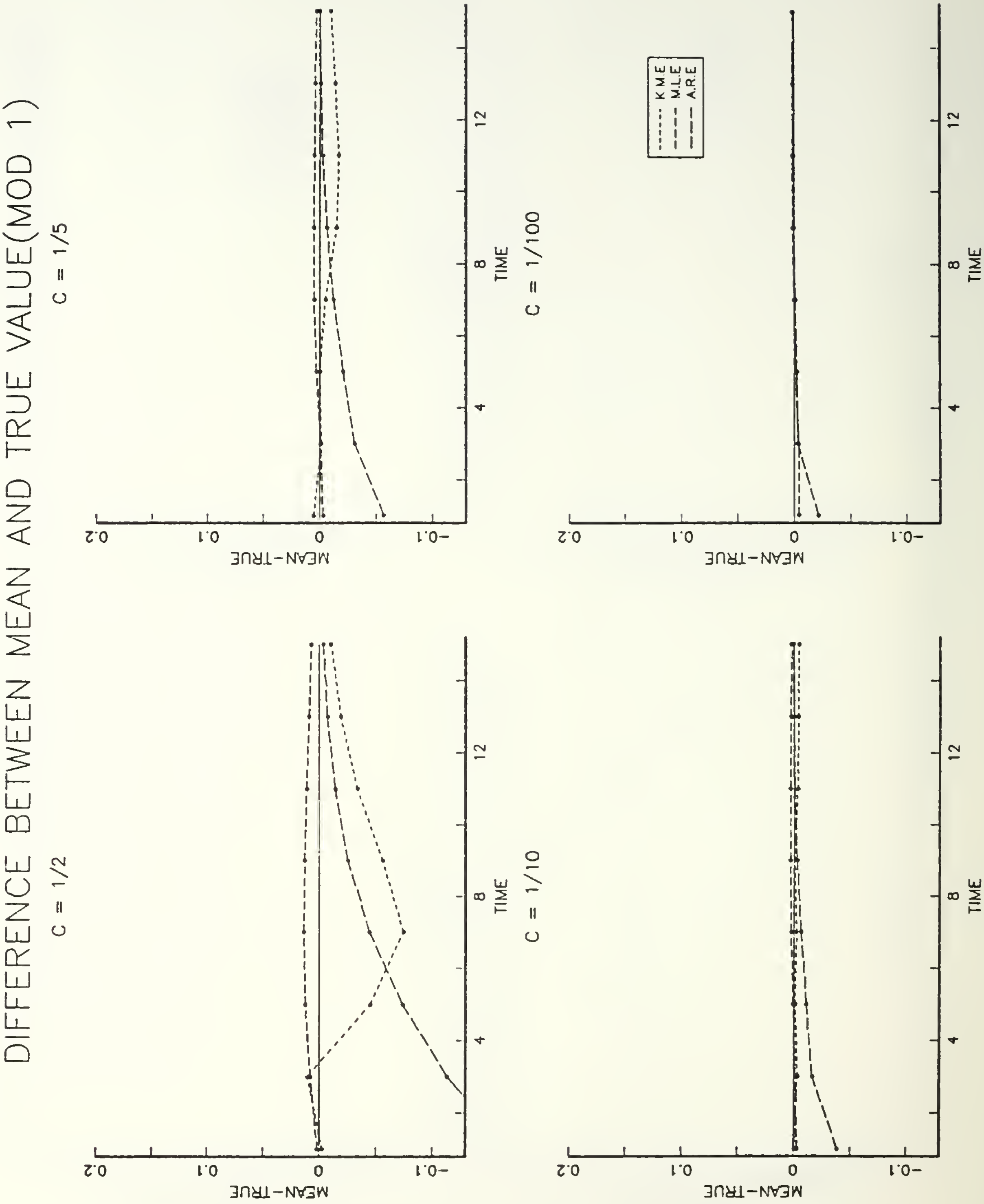

Figure 7. COMPARISON BETWEEN ESTIMATORS(MOD 1): Difference between Mean and True value.

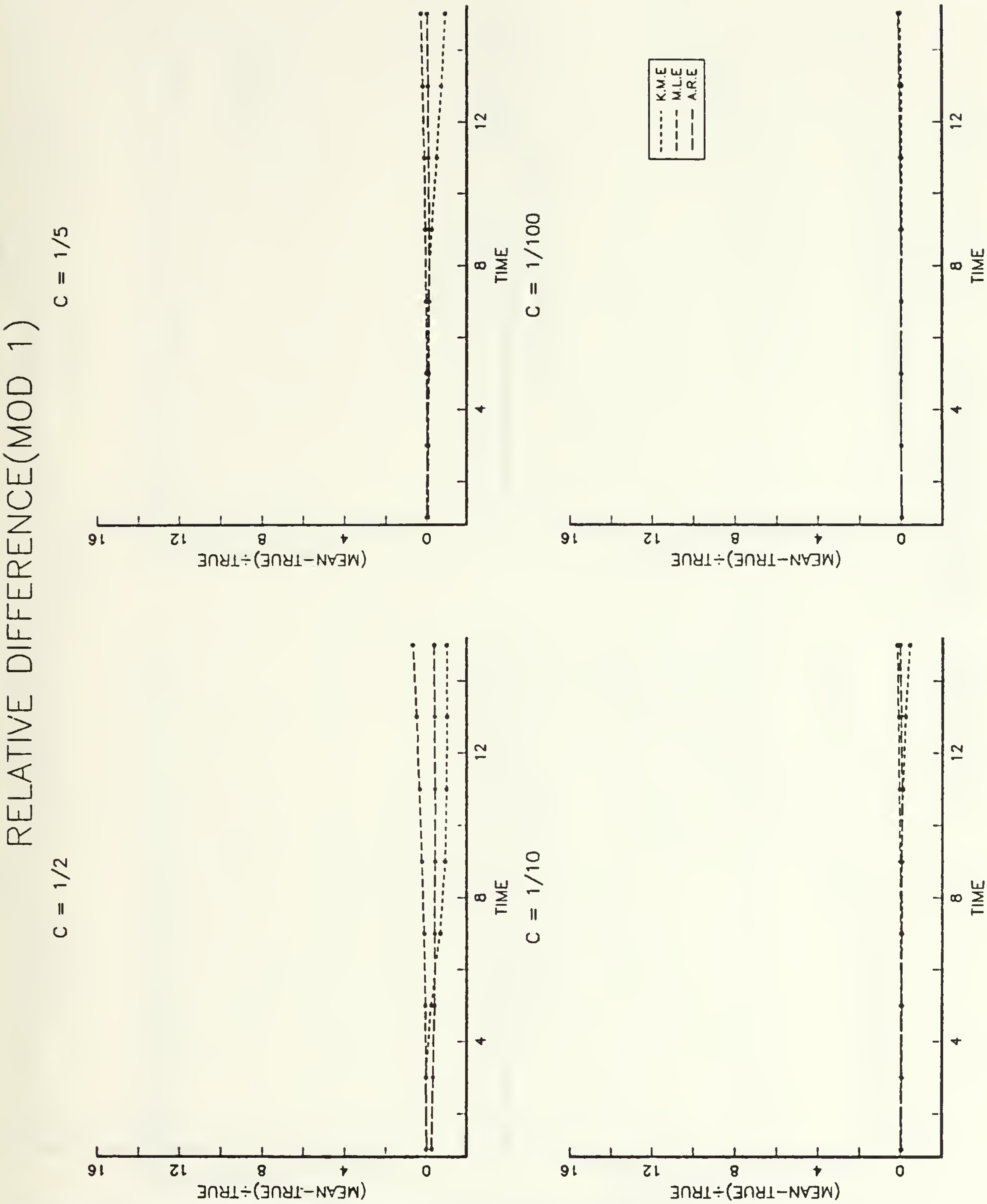

Figure 8. COMPARISON BETWEEN ESTIMATORS(MOD 1): Relative Difference between Mean and True value.

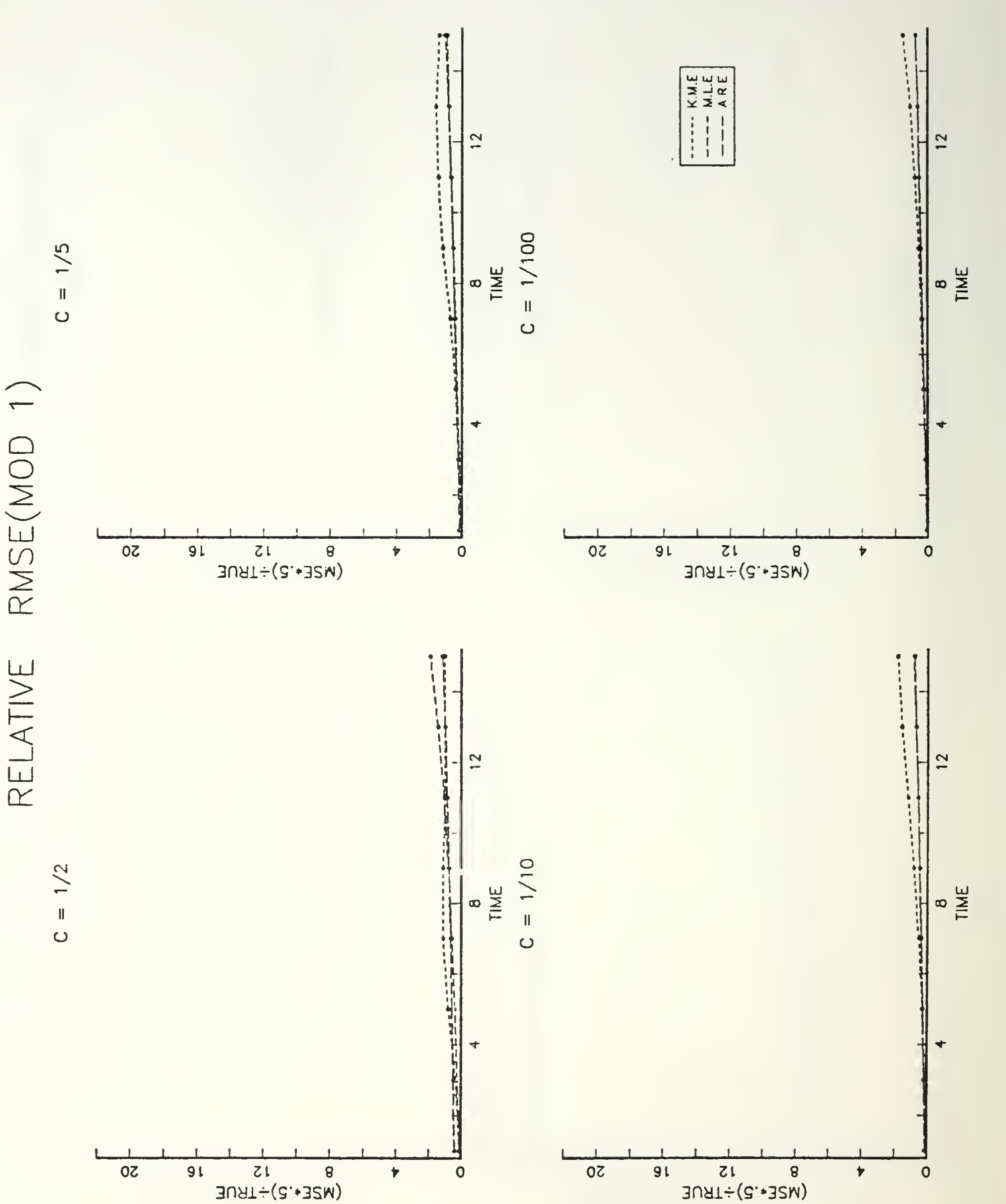

Figure 9. COMPARISON BETWEEN ESTIMATORS(MOD 1): Relative Root Mean Square Error.

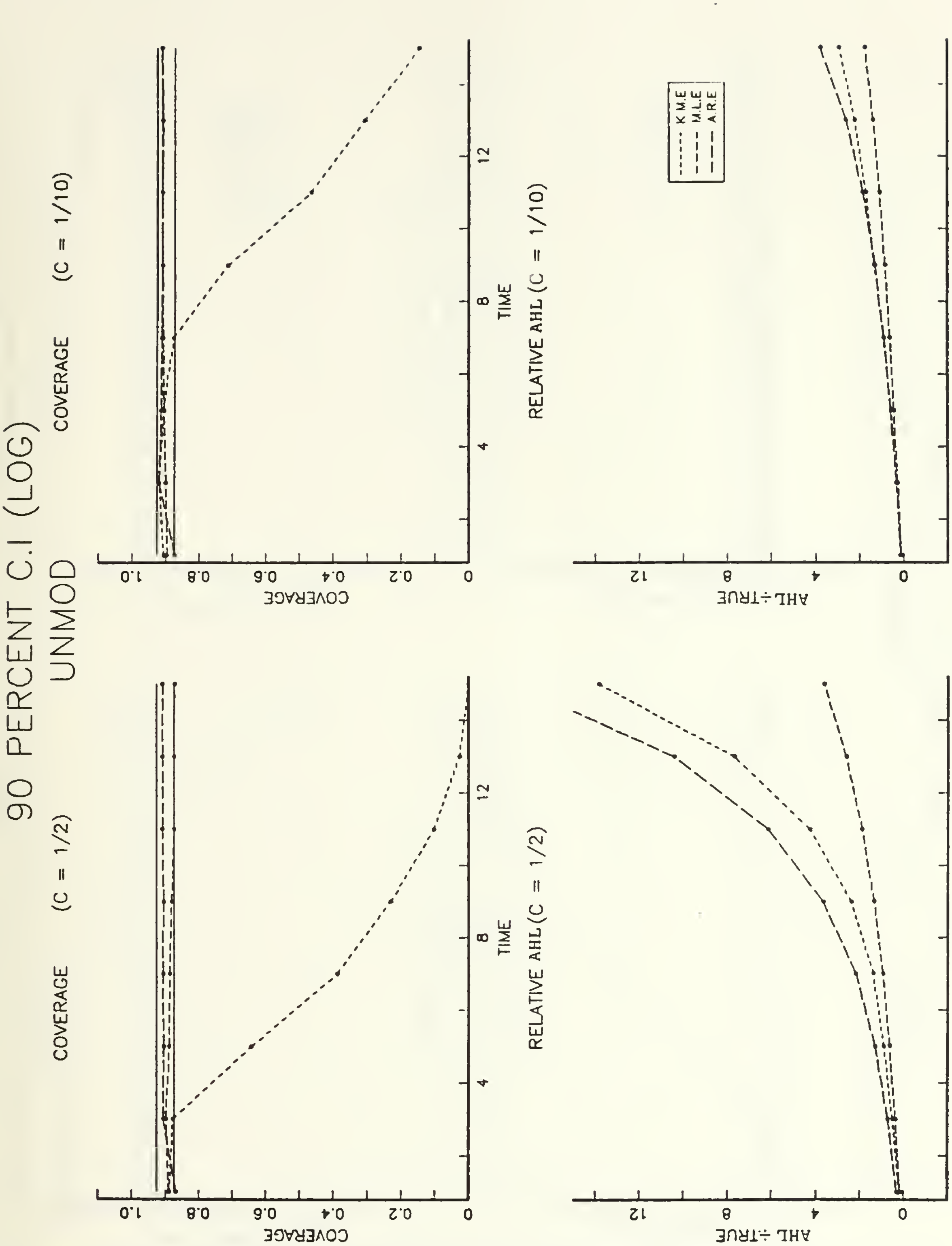

Figure 10. CONFIDENCE INTERVAL(90%, LOG, UNMOD): Coverage Fraction and Average Half Length.

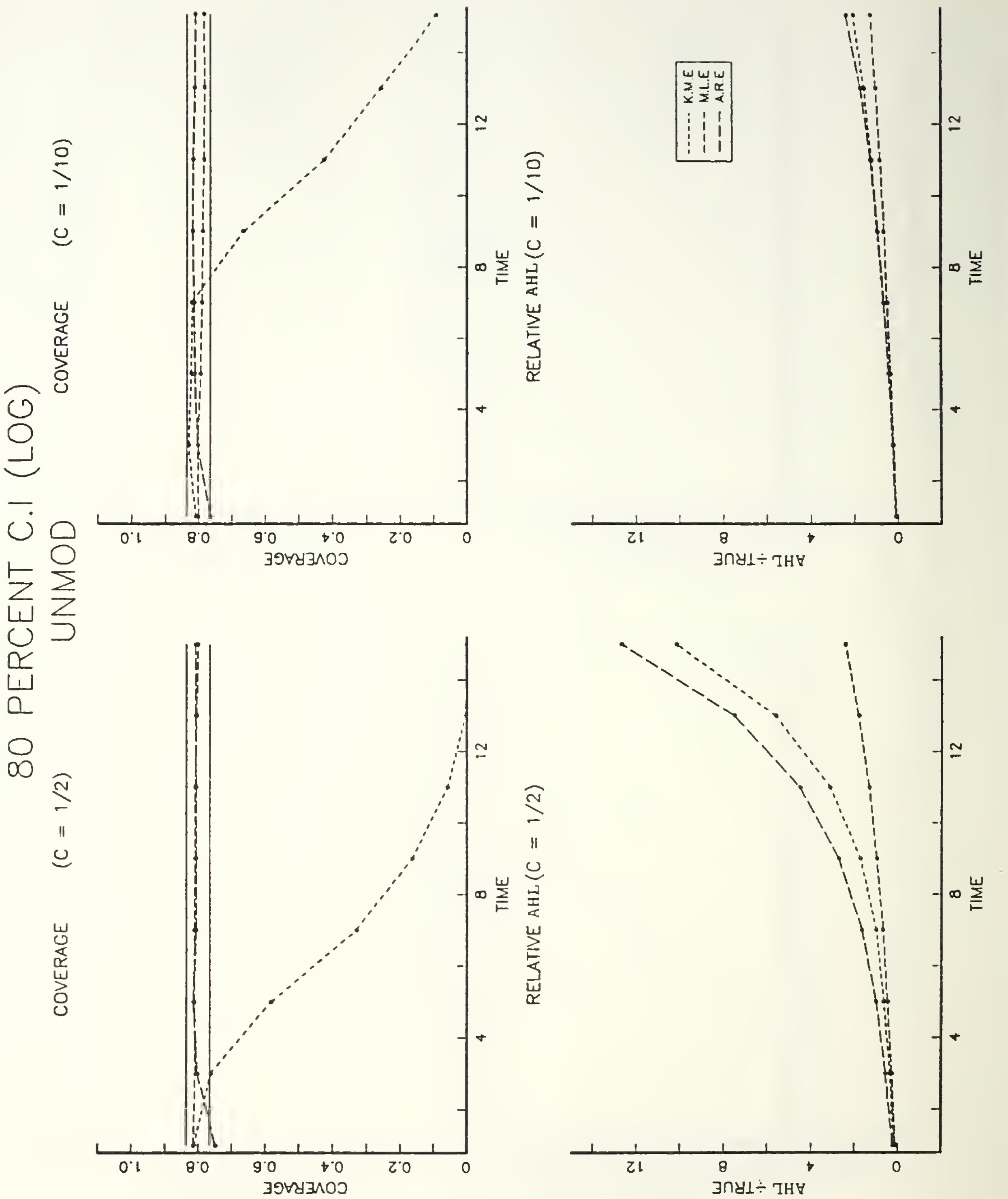

Figure 11. **CONFIDENCE INTERVAL(80%, LOG, UNMOD):** Coverage Fraction and Average Half Length.

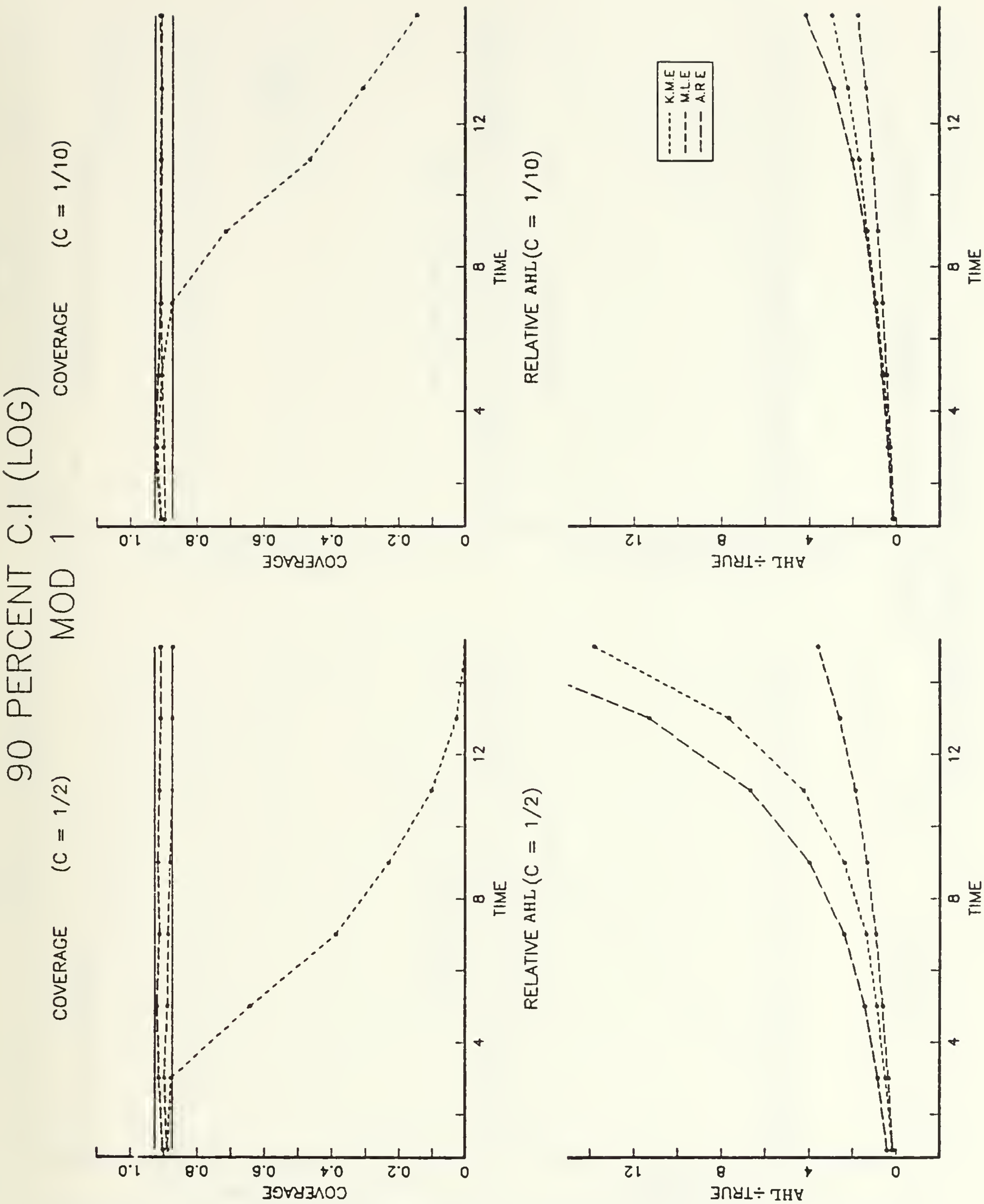

Figure 12. **CONFIDENCE INTERVAL(90%, LOG, MOD 1):** Coverage Fraction and Average Half Length.

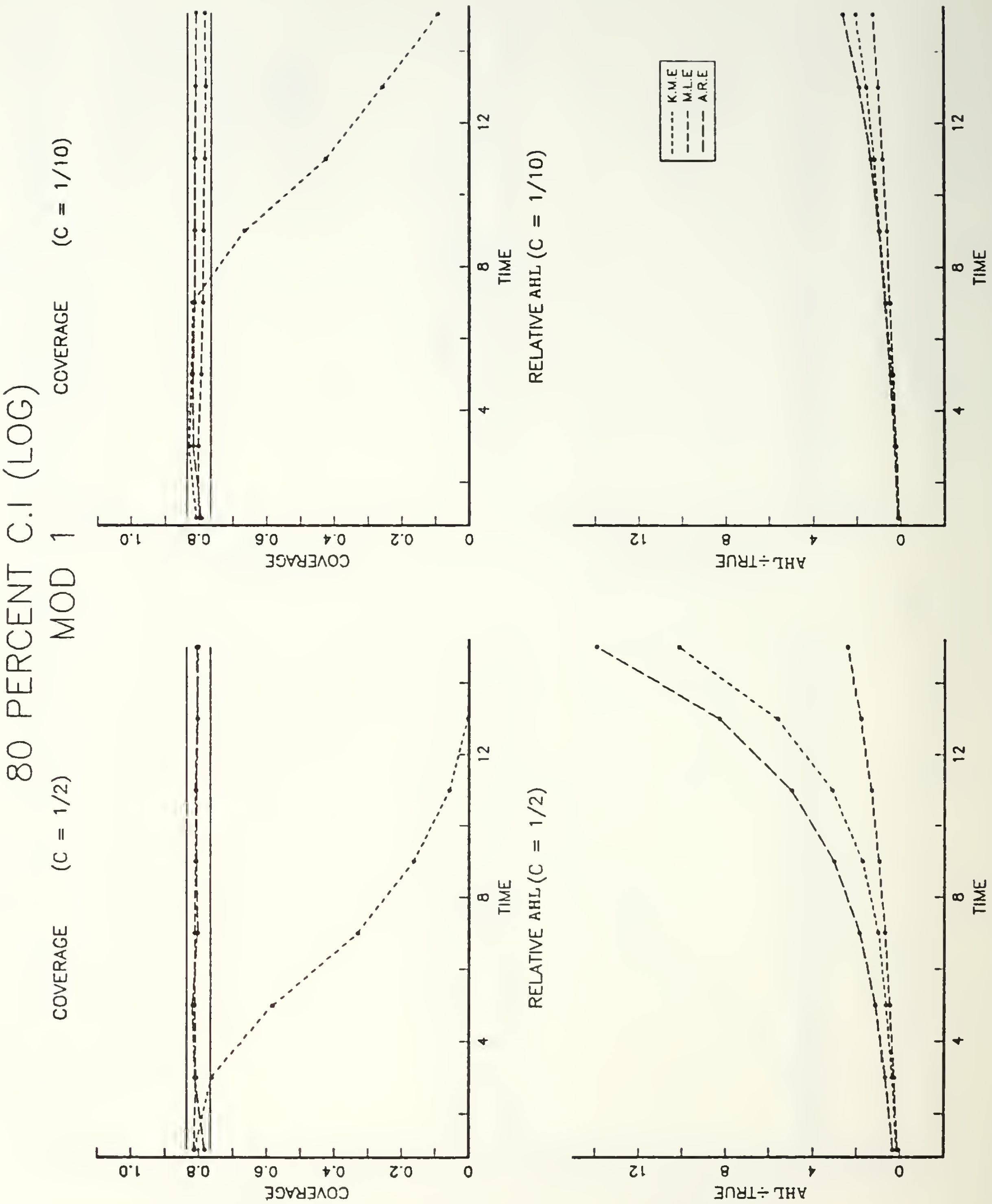

Figure 13. **CONFIDENCE INTERVAL(80%, LOG, MOD 1):** Coverage Fraction and Average Half Length.

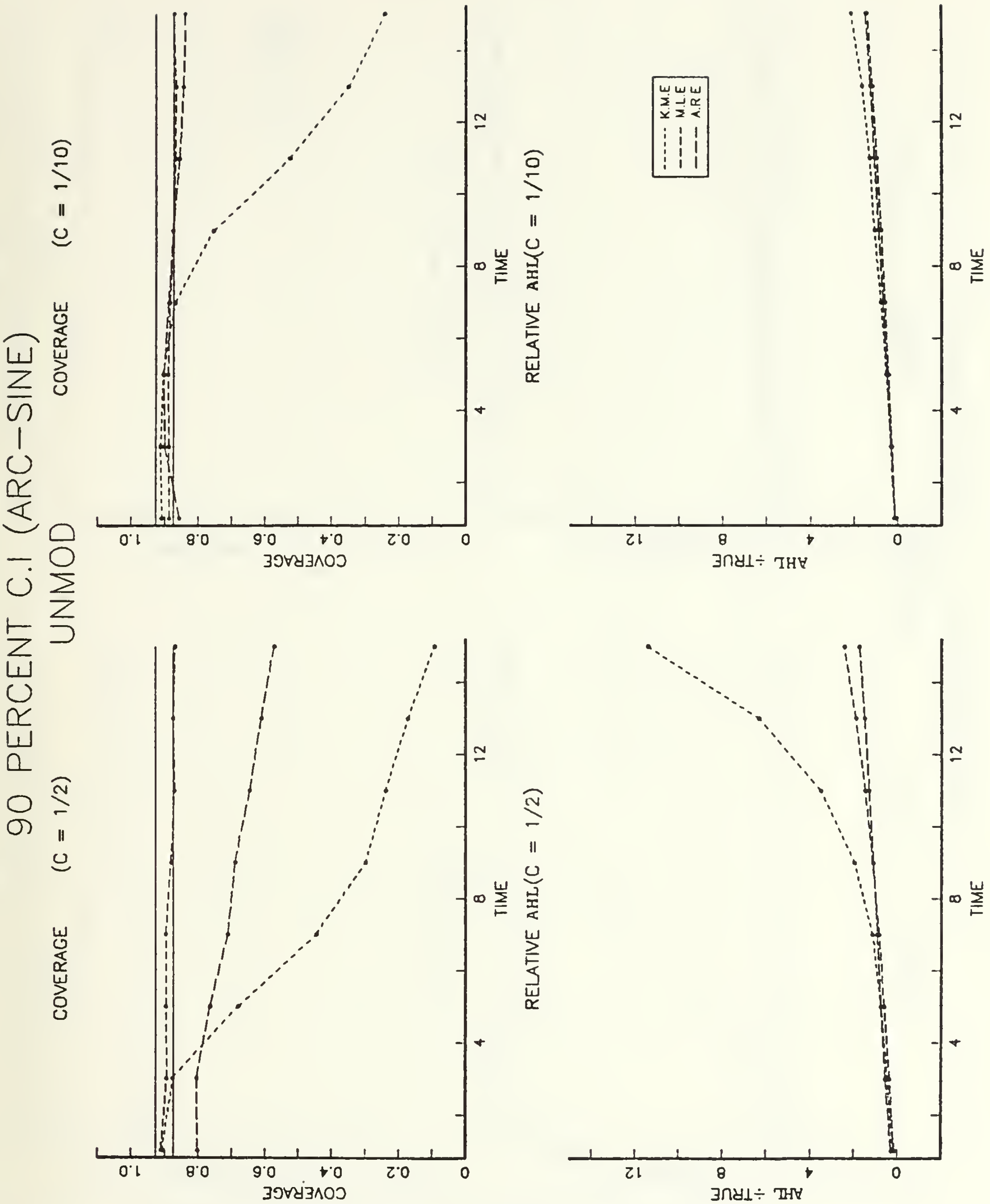

Figure 14. CONFIDENCE INTERVAL(90%, ARC-SINE, UNMOD): Coverage Fraction and Average Half Length.

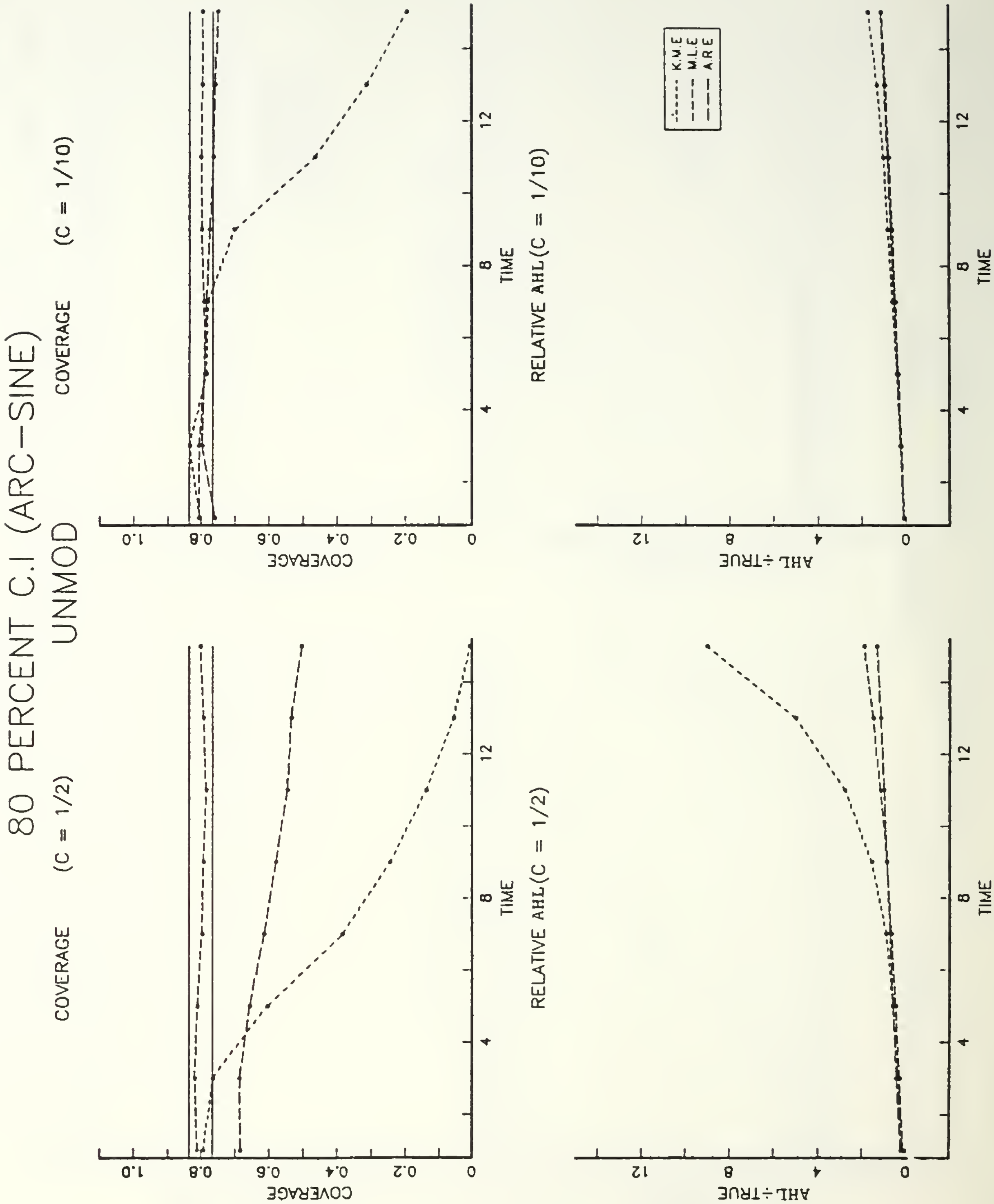

Figure 15. CONFIDENCE INTERVAL(80%, ARC-SINE, UNMOD): Coverage Fraction and Average Half Length.

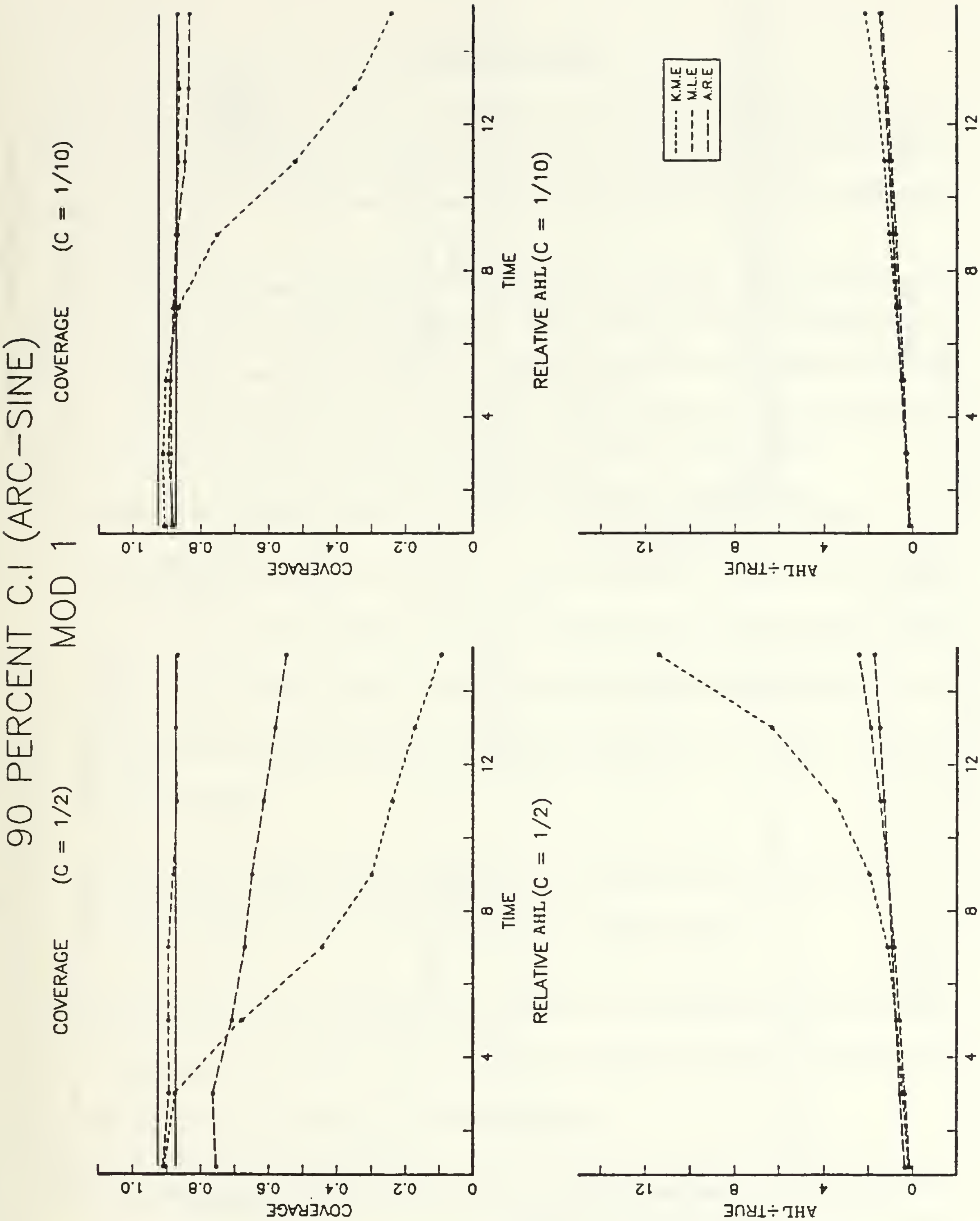

Figure 16. **CONFIDENCE INTERVAL(90%, ARC-SINE, MOD 1):** Coverage Fraction and Average Half Length.

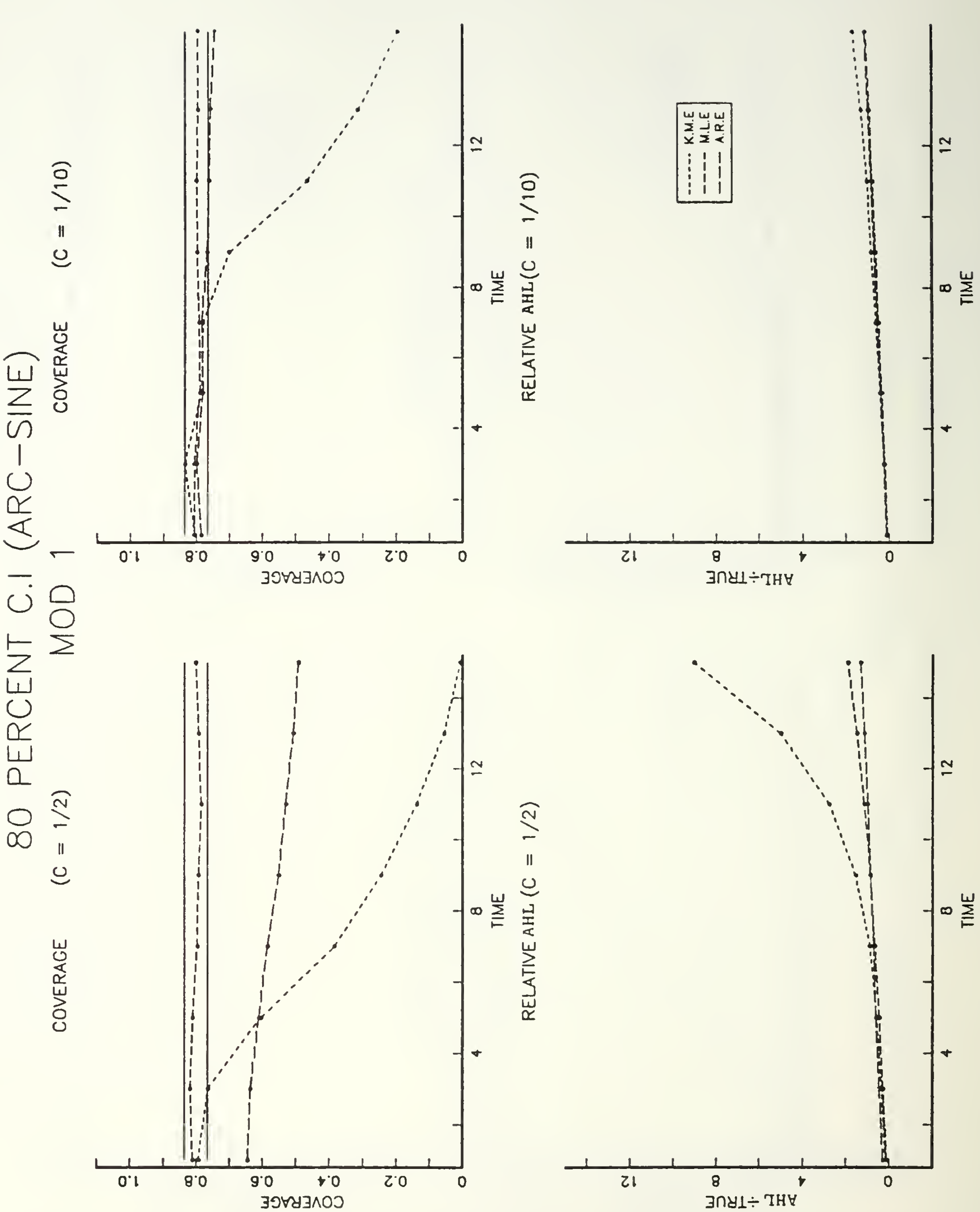

Figure 17. **CONFIDENCE INTERVAL(80%, ARC-SINE, MOD 1):** Coverage Fraction and Average Half Length.

IV. CONCLUSIONS

This thesis considers the problem of estimating the survival probability $P\{D > t\}$ for the first passage time to state 0 for a semi-Markov process using censored data. Simulation is used to study the small sample behavior of three estimators and their confidence interval procedures.

One of the estimators studied is the Kaplan-Meier estimator of the first passage times to state 0. Both the unmodified Kaplan-Meier estimator and two modifications, MOD 1 and MOD 2 making the estimated distribution honest are considered. Another estimator is the maximum likelihood estimator. A third estimator, the asymptotic renewal estimator, uses an exponential approximation to the survival function.

The following conclusions are drawn from the simulation experiment.

1. The modified Kaplan-Meier estimators MOD 1 and MOD 2 using the first passage times to state 0 have a smaller bias for large times than the unmodified K.M.E.. However, in the medium range of times, the two modified procedures MOD 1 and MOD 2 have a larger bias than that of the unmodified K.M.E.. MOD 1 is slightly better than MOD 2.

2. Modifying the Kaplan-Meier estimates of the sojourn time distributions in the asymptotic renewal estimate does not improve its performance.

3. The asymptotic normal confidence intervals for the Kaplan-Meier estimator of the first passage times to state 0 using the arc-sine transformation have a slightly better coverage than those using the log transformation.

4. The confidence intervals using the log transformed estimators are preferred to those using the arc-sine transformed estimators for the maximum likelihood estimator and the asymptotic renewal estimator.

5. The confidence intervals for the jackknifed log transformed asymptotic renewal estimator have the correct coverage for all but the smallest times. This estimator makes no assumptions concerning the parametric form of the sojourn time distributions. It's expected that it will also perform well in cases in which the sojourn time distributions are not exponential.

6. The confidence intervals for the maximum likelihood estimator have the correct coverage, and also have the smallest average half length. However the estimators depend on the parametric form of the sojourn time distributions. If the parametric form is incorrectly specified, the M.L.E. can be quite biased, Kim[Ref. 1].

APPENDIX A. TRUE SURVIVAL PROBABILITY

Table 1. TRUE SURVIVAL PROBABILITY FOR MODEL

TIME	$P\{D > t\}$
1.00	0.6644
3.00	0.3554
5.00	0.1974
7.00	0.1099
9.00	0.0612
11.00	0.0340
13.00	0.0189
15.00	0.0105

APPENDIX B. NUMBER OF K.M.E. DEFINED AT TIME T

Table 2. NUMBER OF KAPLAN-MEIER ESTIMATES DEFINED AT TIME T: C = 1/2

TIME	KM0	KM1	KM2
1.00	500	500	496
3.00	496	428	369
5.00	366	314	334
7.00	215	308	333
9.00	171	308	333
11.00	165	308	333
13.00	164	308	333
15.00	164	308	333

Table 3. NUMBER OF KAPLAN-MEIER ESTIMATES DEFINED AT TIME T: C = 1/5

TIME	KM0	KM1	KM2
1.00	500	500	500
3.00	500	488	463
5.00	496	436	425
7.00	456	421	419
9.00	375	420	419
11.00	330	420	419
13.00	309	420	419
15.00	305	420	419

Table 4. NUMBER OF KAPLAN-MEIER ESTIMATES DEFINED AT TIME T: C = 1/10

TIME	KM0	KM1	KM2
1.00	500	500	500
3.00	500	497	484
5.00	499	462	455
7.00	489	451	451
9.00	466	447	450
11.00	436	447	450
13.00	412	447	450
15.00	399	447	450

Table 5. NUMBER OF KAPLAN-MEIER ESTIMATES DEFINED AT TIME T: C = 1/100

TIME	KM0	KM1	KM2
1.00	500	500	500
3.00	500	500	499
5.00	500	498	494
7.00	500	494	493
9.00	499	494	493
11.00	499	494	493
13.00	498	494	493
15.00	494	494	493

APPENDIX C. STATISTICS OF THREE ESTIMATORS

Table 6. STATISTICS OF THREE ESTIMATORS : UNMOD, C = 1/2

TIME	K.M.E		M.L.E		A.R.E	
	MEAN	SRMSE	MEAN	SRMSE	MEAN	SRMSE
1.0	0.6608	0.0057	0.6654	0.0033	0.5758	0.0196
3.0	0.3672	0.0124	0.3633	0.0064	0.2739	0.0178
5.0	0.2200	0.0198	0.2098	0.0055	0.1370	0.0099
7.0	0.1778	0.0263	0.1237	0.0037	0.0713	0.0045
9.0	0.1700	0.0342	0.0742	0.0023	0.0383	0.0019
11.0	0.1683	0.0403	0.0452	0.0013	0.0212	0.0008
13.0	0.1683	0.0446	0.0279	0.0007	0.0120	0.0003
15.0	0.1683	0.0472	0.0175	0.0004	0.0070	0.0002

Table 7. STATISTICS OF THREE ESTIMATORS : UNMOD, C = 1/5

TIME	K.M.E		M.L.E		A.R.E	
	MEAN	SRMSE	MEAN	SRMSE	MEAN	SRMSE
1.0	0.6688	0.0048	0.6599	0.0024	0.6373	0.0043
3.0	0.3540	0.0068	0.3547	0.0042	0.3389	0.0056
5.0	0.1982	0.0063	0.2004	0.0033	0.1838	0.0040
7.0	0.1135	0.0054	0.1148	0.0020	0.1014	0.0022
9.0	0.0708	0.0050	0.0665	0.0010	0.0568	0.0011
11.0	0.0507	0.0043	0.0388	0.0005	0.0323	0.0005
13.0	0.0435	0.0042	0.0229	0.0003	0.0185	0.0002
15.0	0.0399	0.0042	0.0136	0.0001	0.0108	0.0001

Table 8. STATISTICS OF THREE ESTIMATORS : UNMOD, C = 1/10

TIME	K.M.E		M.L.E		A.R.E	
	MEAN	SRMSE	MEAN	SRMSE	MEAN	SRMSE
1.0	0.6628	0.0046	0.6604	0.0021	0.6428	0.0031
3.0	0.3518	0.0050	0.3535	0.0035	0.3476	0.0040
5.0	0.1962	0.0041	0.1981	0.0026	0.1905	0.0029
7.0	0.1087	0.0032	0.1124	0.0015	0.1057	0.0016
9.0	0.0644	0.0023	0.0643	0.0008	0.0593	0.0008
11.0	0.0380	0.0017	0.0370	0.0004	0.0336	0.0004
13.0	0.0249	0.0013	0.0215	0.0002	0.0192	0.0002
15.0	0.0179	0.0011	0.0126	0.0001	0.0111	0.0001

Table 9. STATISTICS OF THREE ESTIMATORS : UNMOD, C = 1/100

TIME	K.M.E		M.L.E		A.R.E	
	MEAN	SRMSE	MEAN	SRMSE	MEAN	SRMSE
1.0	0.6633	0.0044	0.6591	0.0015	0.6434	0.0025
3.0	0.3558	0.0045	0.3513	0.0027	0.3530	0.0027
5.0	0.1974	0.0034	0.1956	0.0021	0.1954	0.0021
7.0	0.1078	0.0019	0.1100	0.0012	0.1090	0.0012
9.0	0.0593	0.0011	0.0623	0.0006	0.0613	0.0006
11.0	0.0327	0.0007	0.0355	0.0003	0.0347	0.0003
13.0	0.0179	0.0004	0.0204	0.0001	0.0198	0.0001
15.0	0.0110	0.0003	0.0118	0.0001	0.0114	0.0001

Table 10. STATISTICS OF THREE ESTIMATORS : MOD 1, C = 1/2

TIME	K.M.E		M.L.E		A.R.E	
	MEAN	SRMSE	MEAN	SRMSE	MEAN	SRMSE
1.0	0.6608	0.0057	0.6654	0.0033	0.4914	0.0665
3.0	0.3659	0.0131	0.3633	0.0064	0.2419	0.0293
5.0	0.1521	0.0259	0.2098	0.0055	0.1239	0.0127
7.0	0.0355	0.0149	0.1237	0.0037	0.0656	0.0052
9.0	0.0050	0.0047	0.0742	0.0023	0.0358	0.0021
11.0	0.0002	0.0012	0.0452	0.0013	0.0200	0.0008
13.0	0.0000	0.0004	0.0279	0.0007	0.0115	0.0003
15.0	0.0000	0.0001	0.0175	0.0004	0.0067	0.0002

Table 11. STATISTICS OF THREE ESTIMATORS : MOD 1, C = 1/5

TIME	K.M.E		M.L.E		A.R.E	
	MEAN	SRMSE	MEAN	SRMSE	MEAN	SRMSE
1.0	0.6688	0.0048	0.6599	0.0024	0.6068	0.0142
3.0	0.3540	0.0068	0.3547	0.0042	0.3247	0.0086
5.0	0.1973	0.0065	0.2004	0.0033	0.1770	0.0049
7.0	0.1046	0.0063	0.1148	0.0020	0.0981	0.0024
9.0	0.0465	0.0051	0.0665	0.0010	0.0551	0.0011
11.0	0.0174	0.0024	0.0388	0.0005	0.0314	0.0005
13.0	0.0055	0.0009	0.0229	0.0003	0.0181	0.0002
15.0	0.0011	0.0002	0.0136	0.0001	0.0105	0.0001

Table 12. STATISTICS OF THREE ESTIMATORS : MOD 1, C = 1/10

TIME	K.M.E		M.L.E		A.R.E	
	MEAN	SRMSE	MEAN	SRMSE	MEAN	SRMSE
1.0	0.6628	0.0046	0.6604	0.0021	0.6247	0.0086
3.0	0.3518	0.0050	0.3535	0.0035	0.3389	0.0056
5.0	0.1962	0.0041	0.1981	0.0026	0.1862	0.0034
7.0	0.1074	0.0034	0.1124	0.0015	0.1035	0.0018
9.0	0.0605	0.0025	0.0643	0.0008	0.0582	0.0008
11.0	0.0304	0.0016	0.0370	0.0004	0.0330	0.0004
13.0	0.0149	0.0009	0.0215	0.0002	0.0189	0.0002
15.0	0.0060	0.0004	0.0126	0.0001	0.0109	0.0001

Table 13. STATISTICS OF THREE ESTIMATORS : MOD 1, C = 1/100

TIME	K.M.E		M.L.E		A.R.E	
	MEAN	SRMSE	MEAN	SRMSE	MEAN	SRMSE
1.0	0.6633	0.0044	0.6591	0.0015	0.6419	0.0027
3.0	0.3558	0.0045	0.3513	0.0027	0.3522	0.0028
5.0	0.1974	0.0034	0.1956	0.0021	0.1950	0.0021
7.0	0.1078	0.0019	0.1100	0.0012	0.1088	0.0012
9.0	0.0592	0.0011	0.0623	0.0006	0.0612	0.0006
11.0	0.0326	0.0007	0.0355	0.0003	0.0347	0.0003
13.0	0.0178	0.0004	0.0204	0.0001	0.0198	0.0001
15.0	0.0108	0.0003	0.0118	0.0001	0.0114	0.0001

Table 14. STATISTICS OF THREE ESTIMATORS : MOD 2, C = 1/2

TIME	K.M.E		M.L.E		A.R.E	
	MEAN	SRMSE	MEAN	SRMSE	MEAN	SRMSE
1.0	0.6608	0.0057	0.6654	0.0033	0.4966	0.0621
3.0	0.3426	0.0162	0.3633	0.0064	0.2439	0.0284
5.0	0.1050	0.0233	0.2098	0.0055	0.1247	0.0125
7.0	0.0205	0.0116	0.1237	0.0037	0.0659	0.0052
9.0	0.0032	0.0041	0.0742	0.0023	0.0359	0.0021
11.0	0.0002	0.0012	0.0452	0.0013	0.0201	0.0008
13.0	0.0000	0.0004	0.0279	0.0007	0.0115	0.0003
15.0	0.0000	0.0001	0.0175	0.0004	0.0067	0.0002

Table 15. STATISTICS OF THREE ESTIMATORS : MOD 2, C = 1/5

TIME	K.M.E		M.L.E		A.R.E	
	MEAN	SRMSE	MEAN	SRMSE	MEAN	SRMSE
1.0	0.6688	0.0048	0.6599	0.0024	0.6079	0.0135
3.0	0.3540	0.0068	0.3547	0.0042	0.3252	0.0084
5.0	0.1946	0.0068	0.2004	0.0033	0.1772	0.0049
7.0	0.0984	0.0061	0.1148	0.0020	0.0982	0.0024
9.0	0.0420	0.0046	0.0665	0.0010	0.0552	0.0011
11.0	0.0161	0.0022	0.0388	0.0005	0.0314	0.0005
13.0	0.0054	0.0009	0.0229	0.0003	0.0181	0.0002
15.0	0.0009	0.0002	0.0136	0.0001	0.0105	0.0001

Table 16. STATISTICS OF THREE ESTIMATORS : MOD 2, C = 1/10

TIME	K.M.E		M.L.E		A.R.E	
	MEAN	SRMSE	MEAN	SRMSE	MEAN	SRMSE
1.0	0.6628	0.0046	0.6604	0.0021	0.6254	0.0081
3.0	0.3518	0.0050	0.3535	0.0035	0.3392	0.0055
5.0	0.1959	0.0041	0.1981	0.0026	0.1864	0.0034
7.0	0.1067	0.0034	0.1124	0.0015	0.1036	0.0017
9.0	0.0594	0.0024	0.0643	0.0008	0.0582	0.0008
11.0	0.0294	0.0015	0.0370	0.0004	0.0330	0.0004
13.0	0.0140	0.0008	0.0215	0.0002	0.0189	0.0002
15.0	0.0057	0.0003	0.0126	0.0001	0.0109	0.0001

Table 17. STATISTICS OF THREE ESTIMATORS : MOD 2, C = 1/100

TIME	K.M.E		M.L.E		A.R.E	
	MEAN	SRMSE	MEAN	SRMSE	MEAN	SRMSE
1.0	0.6633	0.0044	0.6591	0.0015	0.6419	0.0027
3.0	0.3558	0.0045	0.3513	0.0027	0.3522	0.0028
5.0	0.1974	0.0034	0.1956	0.0021	0.1950	0.0021
7.0	0.1078	0.0019	0.1100	0.0012	0.1088	0.0012
9.0	0.0592	0.0011	0.0623	0.0006	0.0612	0.0006
11.0	0.0326	0.0007	0.0355	0.0003	0.0347	0.0003
13.0	0.0178	0.0004	0.0204	0.0001	0.0198	0.0001
15.0	0.0108	0.0003	0.0118	0.0001	0.0114	0.0001

APPENDIX D. CONFIDENCE INTERVALS

Table 18. TWO-SIDED 90 % COVERAGE FRACTION(UNMOD, C = 1/2)

TIME	COV-ERAGE	K.M.E		M.L.E		A.R.E	
		LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	toohigh	44(.09)	30(.06)	42(.08)	26(.05)	12(.02)	11(.02)
	cover	444(.89)	453(.91)	445(.89)	455(.91)	434(.87)	401(.80)
	too low	12(.02)	17(.03)	13(.03)	19(.04)	54(.11)	88(.18)
3.0	too high	55(.11)	39(.08)	35(.07)	29(.06)	12(.02)	10(.02)
	cover	439(.88)	438(.88)	449(.90)	447(.89)	453(.91)	402(.80)
	too low	6(.01)	23(.05)	16(.03)	24(.05)	35(.07)	88(.18)
5.0	too high	100(.20)	78(.16)	39(.08)	27(.05)	12(.02)	9(.02)
	cover	321(.64)	340(.68)	444(.89)	448(.90)	452(.90)	382(.76)
	too low	79(.16)	82(.16)	17(.03)	25(.05)	36(.07)	109(.22)
7.0	too high	162(.32)	134(.27)	41(.08)	26(.05)	12(.02)	8(.02)
	cover	194(.39)	222(.44)	443(.89)	448(.90)	453(.91)	355(.71)
	too low	144(.29)	144(.29)	16(.03)	26(.05)	35(.07)	137(.27)
9.0	too high	224(.45)	190(.38)	43(.09)	25(.05)	12(.02)	8(.02)
	cover	115(.23)	149(.30)	440(.88)	440(.88)	452(.90)	344(.69)
	too low	161(.32)	161(.32)	17(.03)	35(.07)	36(.07)	148(.30)
11.0	too high	285(.57)	217(.43)	46(.09)	24(.05)	12(.02)	8(.02)
	cover	51(.10)	119(.24)	437(.87)	435(.87)	454(.91)	322(.64)
	too low	164(.33)	164(.33)	17(.03)	41(.08)	34(.07)	170(.34)
13.0	too high	323(.65)	250(.50)	46(.09)	19(.04)	12(.02)	9(.02)
	cover	13(.03)	86(.17)	437(.87)	437(.87)	454(.91)	305(.61)
	too low	164(.33)	164(.33)	17(.03)	44(.09)	34(.07)	186(.37)
15.0	too high	336(.67)	290(.58)	47(.09)	16(.03)	12(.02)	9(.02)
	cover	0(.00)	46(.09)	436(.87)	434(.87)	454(.91)	285(.57)
	too low	164(.33)	164(.33)	17(.03)	50(.10)	34(.07)	206(.41)

Table 19. AVERAGE AND STANDARD DEVIATION OF HALF LENGTH(90% C.I, UNMOD, C = 1/2)

TIME	K.M.E		M.L.E		A.R.E	
	LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	.1245	.1223	.0959	.0949	.2032	.1862
	(.0107)	(.0102)	(.0080)	(.0077)	(.0841)	(.0719)
3.0	.1783	.1647	.1337	.1289	.2583	.1948
	(.0341)	(.0249)	(.0108)	(.0108)	(.1215)	(.0824)
5.0	.1775	.1494	.1251	.1164	.2572	.1490
	(.0904)	(.0722)	(.0223)	(.0223)	(.1506)	(.0840)
7.0	.1535	.1267	.1045	.0927	.2415	.1031
	(.1083)	(.0864)	(.0295)	(.0287)	(.1704)	(.0778)
9.0	.1463	.1207	.0830	.0695	.2243	.0683
	(.1108)	(.0887)	(.0320)	(.0299)	(.1831)	(.0660)
11.0	.1452	.1197	.0644	.0505	.2091	.0444
	(.1113)	(.0891)	(.0316)	(.0281)	(.1914)	(.0531)
13.0	.1452	.1197	.0495	.0361	.1966	.0285
	(.1113)	(.0891)	(.0295)	(.0249)	(.1969)	(.0414)
15.0	.1452	.1197	.0379	.0256	.1866	.0183
	(.1113)	(.0891)	(.0267)	(.0213)	(.2009)	(.0316)

Table 20. TWO-SIDED 80 % COVERAGE FRACTION(UNMOD, C = 1/2)

TIME	COV-ERAGE	K.M.E		M.L.E		A.R.E	
		LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	too high	68(.14)	59(.12)	61(.12)	54(.11)	33(.07)	29(.06)
	cover	406(.81)	396(.79)	407(.81)	405(.81)	374(.75)	342(.68)
	too low	26(.05)	45(.09)	32(.06)	41(.08)	93(.19)	129(.26)
3.0	too high	94(.19)	75(.15)	62(.12)	47(.09)	29(.06)	26(.05)
	cover	381(.76)	382(.76)	404(.80)	409(.82)	401(.80)	343(.69)
	too low	25(.05)	43(.09)	34(.07)	44(.09)	70(.14)	131(.26)
5.0	too high	129(.26)	108(.22)	64(.13)	48(.10)	25(.05)	19(.04)
	cover	291(.58)	302(.60)	406(.81)	405(.81)	407(.81)	328(.76)
	too low	80(.16)	90(.18)	30(.06)	47(.09)	68(.14)	153(.31)
7.0	too high	191(.38)	164(.33)	65(.13)	48(.10)	24(.05)	17(.03)
	cover	165(.33)	192(.38)	406(.81)	398(.80)	403(.81)	307(.61)
	too low	144(.29)	144(.29)	29(.06)	54(.11)	73(.15)	176(.35)
9.0	too high	257(.51)	217(.43)	67(.13)	47(.09)	22(.04)	16(.03)
	cover	82(.16)	122(.24)	405(.81)	396(.79)	403(.81)	289(.58)
	too low	161(.32)	161(.32)	28(.06)	57(.11)	75(.15)	195(.39)
11.0	too high	307(.61)	267(.53)	66(.13)	45(.09)	21(.04)	15(.03)
	cover	29(.06)	69(.14)	405(.81)	392(.78)	403(.81)	273(.55)
	too low	164(.33)	164(.33)	29(.06)	63(.13)	76(.15)	212(.42)
13.0	too high	334(.67)	308(.62)	69(.14)	39(.08)	21(.04)	12(.02)
	cover	2(.00)	28(.06)	402(.80)	396(.79)	404(.81)	267(.53)
	too low	164(.33)	164(.33)	29(.06)	65(.13)	75(.15)	221(.44)
15.0	too high	336(.67)	333(.67)	70(.14)	34(.07)	21(.04)	12(.02)
	cover	0(.00)	3(.01)	400(.80)	400(.80)	404(.81)	252(.50)
	too low	164(.33)	164(.33)	30(.06)	66(.13)	75(.15)	236(.47)

Table 21. AVERAGE AND STANDARD DEVIATION OF HALF LENGTH(80% C.I, UNMOD, C = 1/2)

TIME	K.M.E		M.L.E		A.R.E	
	LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	.0968	.0957	.0746	.0742	.1596	.1447
	(.0082)	(.0080)	(.0062)	(.0061)	(.0748)	(.0602)
3.0	.1359	.1296	.1032	.1010	.2048	.1510
	(.0240)	(.0200)	(.0085)	(.0085)	(.1153)	(.0677)
5.0	.1311	.1181	.0953	.0912	.2010	.1152
	(.0655)	(.0574)	(.0175)	(.0175)	(.14300	(.0677)
7.0	.1126	.1002	.0781	.0726	.1845	.0795
	(.0784)	(.0686)	(.0229)	(.0226)	(.1598)	(.0621)
9.0	.1073	.0955	.0607	.0545	.1677	.0526
	(.0804)	(.0704)	(.0245)	(.0236)	(.1701)	(.0522)
11.0	.1065	.0947	.0459	.0396	.1534	.0340
	(.0808)	(.0707)	(.0237)	(.0221)	(.1765)	(.0417)
13.0	.1065	.0947	.0343	.0283	.1420	.0218
	(.0808)	(.0707)	(.0217)	(.0196)	(.1805)	(.0322)
15.0	.1065	.0947	.0256	.0200	.1330	.0140
	(.0808)	(.0707)	(.0192)	(.0167)	(.1832)	(.0245)

Table 22. TWO-SIDED 90 % COVERAGE FRACTION(UNMOD, C = 1/10)

TIME	COV-ERAGE	K.M.E		M.L.E		A.R.E	
		LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	too high	34(.07)	26(.05)	31(.06)	24(.05)	16(.03)	12(.02)
	cover	453(.91)	454(.91)	448(.90)	443(.89)	436(.87)	428(.86)
	too low	13(.03)	20(.04)	21(.04)	33(.07)	48(.10)	60(.12)
3.0	too high	24(.05)	16(.03)	27(.05)	21(.04)	24(.05)	15(.03)
	cover	460(.92)	456(.91)	450(.90)	444(.89)	460(.92)	450(.90)
	too low	16(.03)	28(.06)	23(.05)	35(.07)	16(.03)	35(.07)
5.0	too high	37(.07)	20(.04)	27(.05)	19(.04)	23(.05)	13(.03)
	cover	453(.91)	452(.90)	452(.90)	445(.89)	457(.91)	451(.90)
	too low	10(.02)	28(.06)	21(.04)	36(.07)	20(.04)	36(.07)
7.0	too high	37(.07)	25(.05)	27(.05)	18(.04)	23(.05)	12(.02)
	cover	438(.88)	434(.87)	454(.91)	442(.88)	456(.91)	444(.89)
	too low	25(.05)	41(.08)	19(.04)	40(.08)	21(.04)	44(.09)
9.0	too high	47(.09)	27(.05)	26(.05)	16(.03)	24(.05)	9(.02)
	cover	357(.71)	377(.75)	455(.91)	438(.88)	455(.91)	438(.88)
	too low	96(.19)	96(.19)	19(.04)	46(.09)	21(.04)	53(.11)
11.0	too high	59(.12)	29(.06)	26(.05)	12(.02)	24(.05)	7(.01)
	cover	233(.47)	263(.53)	455(.91)	434(.87)	455(.91)	428(.86)
	too low	208(.42)	208(.42)	19(.04)	54(.11)	21(.04)	65(.13)
13.0	too high	59(.12)	38(.08)	26(.05)	11(.02)	24(.05)	6(.01)
	cover	157(.31)	175(.35)	454(.91)	433(.87)	455(.91)	422(.84)
	too low	287(.57)	287(.57)	20(.04)	56(.11)	21(.04)	72(.14)
15.0	too high	83(.17)	35(.07)	26(.05)	7(.01)	24(.05)	4(.01)
	cover	73(.15)	121(.24)	456(.91)	436(.87)	455(.91)	419(.84)
	too low	344(.69)	344(.69)	18(.04)	57(.11)	21(.04)	77(.15)

Table 23. AVERAGE AND STANDARD DEVIATION OF HALF LENGTH(90% C.I, UNMOD, C = 1/10)

TIME	K.M.E		M.L.E		A.R.E	
	LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	.1121	.1105	.0718	.0714	.0932	.0903
	(.0064)	(.0062)	(.0061)	(.0060)	(.0317)	(.0279)
3.0	.1223	.1185	.0972	.0953	.1212	.1100
	(.0062)	(.0066)	(.0042)	(.0044)	(.0421)	(.0297)
5.0	.1123	.1046	.0871	.0837	.1160	.0969
	(.0132)	(.0142)	(.0101)	(.0103)	(.0522)	(.0299)
7.0	.0988	.0847	.0688	.0644	.0981	.0738
	(.0285)	(.0267)	(.0134)	(.0135)	(.0576)	(.0280)
9.0	.0806	.0640	.0512	.0464	.0789	.0523
	(.0437)	(.0363)	(.0139)	(.0137)	(.0596)	(.0241)
11.0	.0573	.0432	.0369	.0321	.0624	.0355
	(.0513)	(.0398)	(.0128)	(.0122)	(.0601)	(.0195)
13.0	.0416	.0305	.0261	.0217	.0494	.0235
	(.0506)	(.0380)	(.0109)	(.0101)	(.0601)	(.0151)
15.0	.0306	.0223	.0182	.0145	.0395	.0153
	(.0472)	(.0350)	(.0089)	(.0079)	(.0602)	(.0113)

Table 24. TWO-SIDED 80 % COVERAGE FRACTION(UNMOD, C = 1/10)

TIME	COV-ERAGE	K.M.E		M.L.E		A.R.E	
		LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	too high	61(.12)	47(.09)	56(.11)	48(.10)	38(.08)	31(.06)
	cover	404(.81)	402(.80)	400(.80)	403(.81)	382(.76)	379(.76)
	too low	35(.07)	51(.10)	44(.09)	49(.10)	80(.16)	90(.18)
3.0	too high	49(.10)	38(.08)	52(.10)	42(.08)	57(.11)	43(.09)
	cover	415(.83)	417(.83)	401(.80)	403(.81)	401(.80)	398(.80)
	too low	36(.07)	45(.09)	47(.09)	55(.11)	42(.08)	59(.12)
5.0	too high	56(.11)	47(.09)	52(.10)	40(.08)	53(.11)	40(.08)
	cover	411(.82)	392(.78)	397(.79)	395(.79)	406(.81)	395(.79)
	too low	33(.07)	61(.12)	51(.10)	65(.13)	41(.08)	65(.13)
7.0	too high	55(.11)	43(.09)	51(.10)	39(.08)	50(.01)	35(.07)
	cover	410(.82)	390(.78)	395(.79)	396(.79)	407(.81)	392(.78)
	too low	35(.07)	67(.13	54(.11)	65(.13)	43(.09)	73(.15)
9.0	too high	71(.14)	53(.11)	51(.10)	34(.07)	50(.10)	32(.06)
	cover	333(.67)	351(.70)	394(.79)	399(.80)	409(.82)	387(.77)
	too low	96(.19)	96(.19)	55(.12)	67(.13)	41(.08)	81(.16)
11.0	too high	79(.16)	59(.12)	53(.11)	31(.06)	49(.10)	26(.05)
	cover	213(.43)	233(.47)	392(.78)	400(.80)	408(.82)	382(.76)
	too low	208(.42)	208(.42)	55(.11)	69(.14)	43(.09)	92(.18)
13.0	too high	84(.17)	56(.11)	55(.11)	29(.06)	49(.10)	25(.05)
	cover	129(.26)	157(.31)	391(.78)	398(.89)	406(.81)	379(.76)
	too low	287(.57)	287(.57)	54(.11)	73(.15)	45(.09)	96(.19)
15.0	too high	110(.22)	58(.12)	54(.11)	26(.05)	50(.10)	21(.04)
	cover	46(.09)	98(.20)	392(.78)	398(.80)	405(.81)	375(.75)
	too low	344(.69)	844(.69)	54(.11)	76(.15)	45(.09)	104(.21)

Table 25. AVERAGE AND STANDARD DEVIATION OF HALF LENGTH(80% C.I, UNMOD, C = 1/10)

TIME	K.M.E		M.L.E		A.R.E	
	LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	.0872	.0864	.0559	.0557	.0701	.0683
	(.0049)	(.0048)	(.0047)	(.0047)	(.0238)	(.0213)
3.0	.0945	.0927	.0754	.0745	.0905	.0834
	(.0050)	(.0051)	(.0033)	(.0035)	(.0309)	(.0227)
5.0	.0855	.0819	.0670	.0654	.0851	.0736
	(.0106)	(.0112)	(.0080)	(.0081)	(.0369)	(.0230)
7.0	.0729	.0664	.0524	.0503	.0704	.0560
	(.0216)	(.0209)	(.0105)	(.0106)	(.0407)	(.0214)
9.0	.0578	.0502	.0385	.0362	.0551	.0397
	(.0317)	(.0285)	(.0108)	(.0107)	(.0430)	(.0184)
11.0	.0403	.0339	.0273	.0251	.0422	.0269
	(.0364)	(.0312)	(.0098)	(.0095)	(.0432)	(.0149)
13.0	.0290	.0240	.0190	.0169	.0322	.0178
	(.0355)	(.0298)	(.0083)	(.0079)	(.0430)	(.0115)
15.0	.0213	.0175	.0130	.0113	.0247	.0116
	(.0330)	(.0275)	(.0067)	(.0062)	(.0427)	(.0085)

Table 26. TWO-SIDED 90 % COVERAGE FRACTION(MOD 1, C = 1/2)

TIME	COV-ERAGE	K.M.E		M.L.E		A.R.E	
		LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	too high	44(.09)	30(.06)	42(.08)	26(.05)	9(.02)	8(.02)
	cover	444(.89)	453(.91)	445(.89)	455(.91)	451(.90)	377(.75)
	too low	12(.02)	17(.03)	13(.03)	19(.04)	40(.08)	115(.24)
3.0	too high	55(.11)	39(.08)	35(.07)	29(.06)	9(.02)	7(.01)
	cover	439(.88)	438(.88)	449(.90)	447(.89)	457(.91)	382(.76)
	too low	6(.01)	23(.05)	16(.03)	24(.05)	34(.07)	111(.22)
5.0	too high	100(.20)	78(.16)	39(.08)	27(.05)	9(.02)	10(.02)
	cover	321(.64)	340(.68)	444(.89)	448(.90)	460(.92)	354(.71)
	too low	79(.16)	82(.16)	17(.03)	25(.05)	31(.06)	136(.27)
7.0	too high	162(.32)	134(.27)	41(.08)	26(.05)	10(.02)	10(.02)
	cover	194(.39)	222(.44)	443(.89)	448(.90)	456(.91)	335(.67)
	too low	144(.29)	144(.29)	16(.03)	26(.05)	34(.07)	155(.31)
9.0	too high	224(.45)	190(.38)	43(.09)	25(.05)	10(.02)	10(.02)
	cover	115(.23)	149(.30)	440(.88)	440(.88)	458(.92)	324(.65)
	too low	161(.32)	161(.32)	17(.03)	35(.07)	32(.06)	166(.33)
11.0	too high	285(.57)	217(.43)	46(.09)	24(.05)	11(.02)	10(.02)
	cover	51(.10)	119(.24)	437(.87)	435(.87)	456(.91)	307(.61)
	too low	164(.33)	164(.33)	17(.03)	41(.08)	33(.07)	183(.37)
13.0	too high	323(.65)	250(.50)	46(.09)	19(.04)	11(.02)	11(.02)
	cover	13(.03)	86(.17)	437(.87)	437(.87)	454(.91)	290(.58)
	too low	164(.33)	164(.33)	17(.03)	44(.09)	35(.07)	119(.40)
15.0	too high	336(.67)	290(.58)	47(.09)	16(.03)	11(.02)	12(.02)
	cover	0(.00)	46(.09)	436(.87)	434(.87)	454(.91)	274(.55)
	too low	164(.33)	164(.33)	17(.03)	51(.10)	35(.07)	214(.43)

Table 27. AVERAGE AND STANDARD DEVIATION OF HALF LENGTH(90% C.I, MOD 1, C = 1/2)

TIME	K.M.E		M.L.E		A.R.E	
	LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	.1245	.1223	.0959	.0949	.2871	.2379
	(.0107)	(.0102)	(.0080)	(.0077)	(.1289)	(.1119)
3.0	.1783	.1647	.1337	.1289	.3048	.2088
	(.0341)	(.0249)	(.0108)	(.0108)	(.1359)	(.1049)
5.0	.1775	.1494	.1251	.1164	.2881	.1527
	(.0904)	(.0722)	(.0223)	(.0223)	(.1586)	(.1030)
7.0	.1535	.1267	.1045	.0927	.2657	.1040
	(.1083)	(.0864)	(.0295)	(.0287)	(.1775)	(.0902)
9.0	.1463	.1207	.0830	.0695	.2448	.0684
	(.1108)	(.0887)	(.0320)	(.0299)	(.1902)	(.0736)
9.0	.1452	.1197	.0644	.0505	.2276	.0442
	(.1113)	(.0891)	(.0316)	(.0281)	(.1991)	(.0578)
9.0	.1452	.1197	.0495	.0361	.2137	.0284
	(.1113)	(.0891)	(.0295)	(.0249)	(.2050)	(.0444)
9.0	.1452	.1197	.0379	.0256	.2023	.0182
	(.1113)	(.0891)	(.0267)	(.0213)	(.2089)	(.0336)

Table 28. TWO-SIDED 80 % COVERAGE FRACTION(MOD 1, C = 1/2)

TIME	COV-ERAGE	K.M.E		M.L.E		A.R.E	
		LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	too high	68(.14)	59(.12)	61(.12)	54(.11)	28(.06)	25(.05)
	cover	406(.81)	396(.79)	407(.81)	405(.81)	391(.78)	322(.64)
	too low	26(.05)	45(.09)	32(.06)	41(.08)	81(.16)	153(.31)
3.0	too high	94(.19)	75(.15)	62(.12)	47(.09)	22(.04)	20(.04)
	cover	381(.76)	382(.76)	404(.81)	409(.82)	406(.81)	318(.64)
	too low	25(.05)	43(.09)	34(.07)	44(.09)	72(.14)	162(.32)
5.0	too high	129(.26)	108(.22)	64(.13)	48(.10)	20(.04)	18(.04)
	cover	291(.58)	302(.60)	406(.81)	405(.81)	409(.82)	305(.61)
	too low	80(.16)	90(.18)	30(.06)	47(.09)	71(.14)	177(.35)
7.0	too high	191(.38)	164(.33)	65(.13)	48(.10)	21(.04)	17(.03)
	cover	165(.33)	192(.38)	406(.81)	398(.80)	401(.80)	292(.58)
	too low	144(.29)	144(.29)	29(.06)	54(.11)	78(.16)	191(.38)
9.0	too high	257(.51)	217(.43)	67(.13)	47(.09)	19(.04)	18(.04)
	cover	82(.16)	122(.24)	405(.81)	396(.79)	403(.81)	275(.55)
	too low	161(.32)	161(.32)	28(.06)	57(.11)	78(.16)	207(.41)
11.0	too high	307(.61)	267(.53)	66(.13)	45(.09)	19(.04)	17(.03)
	cover	29(.06)	69(.14)	405(.81)	392(.78)	403(.81)	264(.53)
	too low	164(.33)	164(.33)	29(.06)	63(.13)	78(.16)	219(.44)
13.0	too high	334(.67)	308(.62)	69(.14)	39(.08)	19(.04)	13(.03)
	cover	2(.00)	28(.06)	402(.80)	396(.79)	401(.80)	253(.51)
	too low	164(.33)	164(.33)	29(.06)	65(.13)	80(.16)	234(.47)
15.0	too high	336(.67)	333(.67)	70(.14)	34(.07)	19(.04)	12(.02)
	cover	0(.00)	3 (.01)	400(.80)	400(.80)	403(.81)	245(.49)
	too low	164(.33)	164(.33)	30(.06)	66(.13)	78(.16)	243(.49)

Table 29. AVERAGE AND STANDARD DEVIATION OF HALF LENGTH(80% C.I, MOD 1, C = 1/2)

TIME	K.M.E		M.L.E		A.R.E	
	LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	.0968	.0957	.0746	.0742	.2321	.1925
	(.0082)	(.0080)	(.0062)	(.0061)	(.1207)	(.0972)
3.0	.1359	.1296	.1032	.1010	.2440	.1648
	(.0240)	(.0200)	(.0085)	(.0085)	(.1349)	(.0879)
5.0	.1311	.1181	.0953	.0912	.2284	.1193
	(.0655)	(.0574)	(.0175)	(.0175)	(.1592)	(.0843)
7.0	.1126	.1002	.0781	.0726	.2069	.0807
	(.0784)	(.0686)	(.0229)	(.0226)	(.1748)	(.0726)
9.0	.1073	.0955	.0607	.0545	.1865	.0528
	(.0804)	(.0704)	(.0245)	(.0236)	(.1829)	(.0586)
11.0	.1065	.0947	.0459	.0396	.1697	.0340
	(.0808)	(.0707)	(.0237)	(.0221)	(.1876)	(.0456)
13.0	.1065	.0947	.0343	.0283	.1566	.0218
	(.0808)	(.0707)	(.0217)	(.0196)	(.1906)	(.0347)
15.0	.1065	.0947	.0256	.0200	.1463	.0139
	(.0808)	(.0707)	(.0192)	(.0167)	(.1924)	(.0261)

Table 30. TWO-SIDED 90 % COVERAGE FRACTION(MOD 1, C = 1/10)

TIME	COVERAGE	K.M.E		M.L.E		A.R.E	
		LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	too high	34(.07)	26(.05)	31(.06)	24(.05)	14(.03)	10(.02)
	cover	453(.91)	454(.91)	448(.90)	443(.89)	454(.91)	440(.88)
	too low	13(.03)	20(.04)	21(.04)	33(.07)	32(.06)	50(.10)
3.0	too high	24(.05)	16(.03)	27(.05)	21(.04)	20(.04)	13(.03)
	cover	460(.92)	456(.91)	450(.90)	444(.89)	463(.93)	448(.90)
	too low	16(.03)	28(.06)	23(.05)	35(.67)	17(.03)	39(.08)
5.0	too high	37(.07)	20(.04)	27(.05)	19(.04)	19(.04)	11(.02)
	cover	453(.91)	452(.90)	452(.90)	445(.89)	458(.92)	446(.89)
	too low	10(.02)	28(.06)	21(.04)	36(.07)	23(.05)	43(.09)
7.0	too high	37(.07)	25(.05)	27(.05)	18(.04)	21(.04)	10(.02)
	cover	438(.88)	434(.87)	454(.91)	442(.88)	455(.91)	439(.88)
	too low	25(.05)	41(.08)	19(.04)	40(.08)	24(.05)	51(.10)
9.0	too high	47(.09)	27(.05)	26(.05)	16(.03)	23(.05)	8(.02)
	cover	357(.71)	377(.75)	455(.91)	438(.88)	454(.91)	435(.87)
	too low	96(.19)	96(.19)	19(.04)	46(.09)	23(.05)	57(.11)
11.0	too high	59(.12)	29(.06)	26(.05)	12(.02)	23(.05)	6(.01)
	cover	233(.47)	263(.53)	455(.91)	434(.87)	453(.91)	424(.85)
	too low	208(.42)	208(.42)	19(.04)	54(.11)	24(.05)	70(.14)
13.0	too high	59(.12)	38(.08)	26(.05)	11(.02)	23(.05)	5(.01)
	cover	154(.31)	175(.35)	454(.91)	433(.87)	453(.91)	419(.84)
	too low	287(.57)	287(.57)	20(.04)	56(.11)	24(.05)	76(.15)
15.0	too high	83(.17)	35(.07)	26(.05)	7(.01)	23(.05)	3(.01)
	cover	73(.15)	121(.24)	456(.91)	436(.87)	453(.91)	418(.84)
	too low	344(.89)	344(.89)	18(.04)	57(.11)	24(.05)	79(.16)

Table 31. AVERAGE AND STANDARD DEVIATION OF HALF LENGTH(90% C.I, MOD 1, C=1/10)

TIME	K.M.E		M.L.E		A.R.E	
	LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	.1121	.1105	.0718	.0714	.1276	.1187
	(.0064)	(.0062)	(.0061)	(.0060)	(.0812)	(.0704)
3.0	.1223	.1185	.0972	.0953	.1431	.1222
	(.0062)	(.0066)	(.0042)	(.0044)	(.0762)	(.0457)
5.0	.1123	.1046	.0871	.0837	.1300	.1021
	(.0132)	(.0142)	(.0101)	(.0103)	(.0765)	(.0379)
7.0	.0988	.0847	.0688	.0644	.1077	.0761
	(.0285)	(.0267)	(.0134)	(.0135)	(.0760)	(.0325)
9.0	.0806	.0640	.0512	.0464	.0861	.0533
	(.0437)	(.0363)	(.0139)	(.0137)	(.0749)	(.0268)
11.0	.0573	.0432	.0369	.0321	.0681	.0360
	(.0513)	(.0398)	(.0128)	(.0122)	(.0733)	(.0211)
13.0	.0416	.0305	.0261	.0217	.0540	.0237
	(.0506)	(.0380)	(.0109)	(.0101)	(.0716)	(.0160)
15.0	.0306	.0223	.0182	.0145	.0433	.0154
	(.0472)	(.0350)	(.0089)	(.0079)	(.0701)	(.0118)

Table 32. TWO-SIDED 80 % COVERAGE FRACTION(MOD 1, C = 1/10)

TIME	COVERAGE	K.M.E		M.L.E		A.R.E	
		LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	too high	61(.12)	47(.09)	56(.11)	48(.10)	35(.07)	28(.06)
	cover	404(.81)	402(.80)	400(.80)	403(.81)	396(.79)	392(.78)
	too low	35(.07)	51(.10)	44(.09)	49(.10)	69(.14)	80(.16)
3.0	too high	49(.10)	38(.08)	52(.10)	42(.08)	49(.10)	37(.07)
	cover	415(.83)	417(.83)	401(.80)	403(.81)	408(.82)	399(.80)
	too low	36(.07)	45(.09)	47(.09)	55(.11)	43(.09)	64(.13)
5.0	too high	56(.11)	47(.09)	52(.10)	40(.08)	49(.10)	39(.08)
	cover	411(.82)	392(.78)	397(.79)	395(.79)	409(.82)	390(.78)
	too low	33(.07)	61(.12)	51(.10)	65(.13)	42(.08)	71(.14)
7.0	too high	55(.11)	43(.09)	51(.10)	39(.08)	47(.09)	34(.07)
	cover	410(.82)	390(.78)	395(.79)	396(.79)	407(.81)	392(.78)
	too low	35(.07)	67(.13)	54(.11)	65(.13)	46(.09)	74(.15)
9.0	too high	71(.14)	53(.11)	51(.10)	34(.07)	48(.10)	31(.06)
	cover	333(.67)	351(.70)	394(.79)	399(.80)	407(.81)	384(.77)
	too low	96(.19)	96(.19)	55(.11)	67(.13)	45(.09)	85(.17)
11.0	too high	79(.16)	59(.12)	53(.11)	31(.06)	47(.09)	25(.05)
	cover	213(.43)	233(.47)	392(.78)	400(.80)	407(.81)	381(.76)
	too low	208(.42)	208(.42)	55(.11)	69(.14)	46(.09)	94(.19)
13.0	too high	84(.17)	56(.11)	55(.11)	29(.06)	47(.09)	24(.05)
	cover	129(.26)	157(.31)	391(.78)	398(.80)	406(.81)	379(.76)
	too low	287(.57)	287(.57)	54(.11)	73(.15)	47(.09)	97(.19)
15.0	too high	110(.22)	58(.12)	54(.11)	26(.05)	48(.10)	21(.04)
	cover	46(.09)	98(.20)	392(.78)	398(.80)	405(.81)	373(.75)
	too low	344(.69)	344(.69)	54(.11)	76(.15)	47(.09)	106(.21)

Table 33. AVERAGE AND STANDARD DEVIATION OF HALF LENGTH(80% C.I, MOD 1, C = 1/10)

TIME	K.M.E		M.L.E		A.R.E	
	LOG	ASIN	LOG	ASIN	LOG	ASIN
1.0	.0872	.0864	.0559	.0557	.0981	.0908
	(.0049)	(.0048)	(.0047)	(.0047)	(.0665)	(.0563)
3.0	.0945	.0927	.0754	.0745	.1070	.0929
	(.0050)	(.0051)	(.0033)	(.0035)	(.0593)	(.0357)
5.0	.0855	.0819	.0670	.0654	.0952	.0776
	(.0106)	(.0112)	(.0080)	(.0081)	(.0574)	(.0293)
7.0	.0729	.0664	.0524	.0503	.0771	.0578
	(.0216)	(.0209)	(.0105)	(.0106)	(.0571)	(.0250)
9.0	.0578	.0502	.0385	.0362	.0600	.0404
	(.0317)	(.0285)	(.0108)	(.0107)	(.0561)	(.0205)
11.0	.0403	.0339	.0273	.0251	.0461	.0273
	(.0364)	(.0312)	(.0098)	(.0095)	(.0548)	(.0161)
13.0	.0290	.0240	.0190	.0169	.0353	.0180
	(.0355)	(.0298)	(.0083)	(.0079)	(.0535)	(.0122)
15.0	.0213	.0175	.0130	.0113	.0274	.0116
	(.0330)	(.0275)	(.0067)	(.0062)	(.0524)	(.0089)

LIST OF REFERENCES

1. Kim, S. W.(1987), A simulation Study of Estimates of a First Passage Time Distribution for a Semi-Markov Process, Master's Thesis, Naval Postgraduate School, Monterey, California.

2. Jacobs, P. A.(1987), Estimation of the Probability of a Long Time to the First Entrance to a State in a Semi-Markov Model, Naval Postgraduate School Technical Report NPS 55-87-012, Monterey, California.

3. Gallagher, R. M.(1986), A simulation Study of Estimate of a First Passage Time Distribution for a Censored Semi-Markov Process, Master's Thesis, Naval Postgraduate School, Monterey, California.

4. Kaplan, E. L. and Meier, P.(1958), Nonparametric Estimation from Incomplete Observations, *J. of the American Statistical Association*, v.53, pp457-481.

5. Jacobs, P. A.(1986), Results of a Simulation Study of Estimates of a First Passage Time Distribution for Censored Semi-Markov Process, Naval Postgraduate School, Monterey, California.

6. Feller, W.(1966), An Introduction to Probability Theory and Its Applications, v.2, John Wiley, New York. p376.

7. Cressie, N.(1981), Transformations and the Jackknife, *J. of Royal Stat. Soc.*, (B), v.43, No.2, pp177-182.

8. Gaver, D. P. and Miller, R. G. Jr.(1983), Jackknifing the Kaplan-Meier Survival Estimator for Censored Data: Simulation Results and Asymtotic Analysis, *Comm. Stat. Theor. Meth.*, v.12, pp1701-1718.

9. Quenouille, M. H.(1956), Notes on Bias in Estimation, *Biometrika*, v.43, pp353-360.

10. Tukey, J. W.(1958), Bias and Confidence in Not Quite Large Samples, *Ann. Math. Stat.,* v.29, p614.

11. Miller, R. G.(1974), The Jackknife - a Review, *Biometrika,* v.61, pp1-15.

12. Mosteller, F. and Tukey, J. W.(1977), Data Analysis and Regression, Addison Wesley, Inc. p162.

13. Lewis, A. W. and Uribe, L.(1981), The New Naval Postgradute School Random Number Package - LLRANDOM II, Naval Postgradute School Technical Report, NPS55-81-005, Monterey California.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Technical Information Center Cameron Station Alexandria, VA 22304-6145	2
2.	Library, Code 0142 Naval Postgraduate School Monterey, CA 93943-5002	2
3.	Professor P. A. Jacobs, Code 55JC Department of Operations Research Naval Postgraduate School Monterey, CA 93943-5000	1
4.	Professor D. P. Gaver, Code 55GV Department of Operations Research Naval Postgraduate School Monterey, CA 93943-5000	1
5.	Professor D. A. Schrady, Code 55SO Department of Operations Research Naval Postgraduate School Monterey, CA 93943-5000	1
6.	Professor Peter Purdue, Code 55PD Department of Operations Research Naval Postgraduate School Monterey, CA 93943-5000	1
7.	Professor Keebom Kang, Code 55KA Department of Operations Research Naval Postgraduate School Monterey, CA 93943-5000	1
8.	Air Force Central Library Postal Code 151-00 Sindaebang Dong, Kwanak Ku, Seoul Republic of Korea	2
9.	Library of Air Force Academy Postal Code 370-72 Chongwon Gun, Chung Cheong Bug Do, Republic of Korea	2

10. Park, Byung Goo 2
Postal Code 130-12
229 Dae Ssang Ryung-Ri,
Cho Wol-Myun, Kwang Ju-Kun, Kyung Ki-Do
Republic of Korea

11. Park, Byung Hee 1
Postal Code 100-00
Sindang-Dong 432-119, 12/5
Chung-Gu Seoul,
Republic of Korea

12. Song, Chong Chul 1
Postal Code 548-880
Won Sang Dae-Burak, Sang Dae-Ri,
Po Du-Moeyn, Ko Heong-Kun, Cheon Nam,
Republic of Korea

13. Lee, Jae Yeong 1
Postal Code 422-011
Simgok-1-Dong 585, 10/4
Kyunggi-Do Bucheon City,
Republic of Korea

14. Lee, Chong Ho 1
Postal Code 215-800
Nammun 1-Ri 3-Ban, Yang Yang-Eup,
Yang Yang-Kun, Kang Won-Do,
Republic of Korea

15. Lt. Newton R. LIMA 1
Brazilian Naval Commission
4706, Wisconsin Avenue, N.W.
Washington D.C. 20016

16. Lee, Chong Ho 1
Postal Code 215-800
Nammun 1-Ri 3-Ban, Yang Yang-Eup,
Yang Yang-Kun, Kang Won-Do,
Republic of Korea

17. Bae, Deuk Hwan 1
SMC #1240
Naval Postgraduate School
Monterey, CA 93943

18. Lee, Jun 1
SMC #1215
Naval Postgraduate School
Monterey, CA 93943

Thesis
P15672 Park
c.1 Simulation study of estimators for the survival probability of a first passage time for a semi-Markov process using censored data.